EQUILIBRIUM STATES FOR INTERVAL MAPS: THE POTENTIAL $-\mathbf{t}\log|\mathbf{Df}|$

HENK BRUIN, MIKE TODD

ABSTRACT. Let $f : I \to I$ be a C^2 multimodal interval map satisfying polynomial growth of the derivatives along critical orbits. We prove the existence and uniqueness of equilibrium states for the potential $\varphi_t : x \mapsto -t\log|Df(x)|$ for t close to 1, and also that the pressure function $t \mapsto P(\varphi_t)$ is analytic on an appropriate interval near $t = 1$.

1. INTRODUCTION

Thermodynamic formalism ties potential functions φ to invariant measures of a dynamical system (X, f). The aim is to identify and prove uniqueness of a measure μ_φ that maximises the *free energy*, i.e., the sum of the entropy and the integral over the potential. In other words

$$h_{\mu_\varphi}(f) + \int_X \varphi \, d\mu_\varphi = P(\varphi) := \sup_{\nu \in \mathcal{M}_{erg}} \left\{ h_\nu(f) + \int_X \varphi \, d\nu : -\int_X \varphi \, d\nu < \infty \right\}$$

where $\mathcal{M}_{erg}$ is the set of all ergodic f-invariant Borel probability measures. Such measures are called *equilibrium states*, and $P(\varphi)$ is the *pressure*. This theory was developed by Sinai, Ruelle and Bowen [Si, Bo, Ru2] in the context of Hölder potentials on hyperbolic dynamical systems, and has been applied to Axiom A systems, Anosov diffeomorphisms and other systems too, see e.g. [Ba, K2] for more recent expositions. Apart from uniqueness, it was shown in this context that the density $\frac{d\mu_\varphi}{dm_\varphi}$ of the invariant measure with respect to φ-conformal measure m_φ is a fixed point of the transfer operator $(\mathcal{L}_\varphi h)(x) = \sum_{f(y)=x} e^{\varphi(y)} \, h(y)$. Moreover, μ_φ is a Gibbs measure, i.e., there is a constant $K > 0$ such that

$$\frac{1}{K} \leqslant \frac{\mu_\varphi(\mathbf{C}_n)}{e^{\varphi_n(x) - nP(\varphi)}} \leqslant K$$

for all $n \in \mathbb{N}$, all n-cylinder sets $\mathbf{C}_n$ and any $x \in \mathbf{C}_n$. Here $\varphi_n(x) := \varphi(f^{n-1}(x)) + \cdots + \varphi(x)$.

In this paper we are interested in interval maps (I, f) with nonempty set Crit of critical points. These maps are, at best, only non-uniformly hyperbolic. We say that c is a non-flat critical point of f if there exists a diffeomorphism $g_c : \mathbb{R} \to \mathbb{R}$

2000 *Mathematics Subject Classification.* 37D35, 37D25, 37E05.

Key words and phrases. Equilibrium states, thermodynamic formalism, interval maps, non-uniform hyperbolicity.

This research was supported by EPSRC grant GR/S91147/01.

with $g_c(0) = 0$ and $1 < \ell_c < \infty$ such that for x close to c, $f(x) = f(c) \pm |\varphi_c(x-c)|^{\ell_c}$. The value of ℓ_c is known as the *critical order* of c. Let $\ell_{max} = \max\{\ell_c : c \in \text{Crit}\}$. Throughout, $\mathcal{H}$ will be the collection of C^2 interval maps with finitely many branches and only non-flat critical points. There is a finite partition $\mathcal{P}_1$ into maximal intervals on which f is monotone. Let us call this partition the *branch partition.* We will assume throughout that $\vee_n \mathcal{P}_n$ generates the Borel σ-algebra. Note that if $f \in \mathcal{H}$ is C^2 and has no attracting cycles then $\vee_n \mathcal{P}_n$ generates the Borel σ-algebra, see [MSt]. (The C^2 assumption precludes wandering sets, which are not very interesting from the measure theoretic point of view anyway.)

The principal examples of maps in $\mathcal{H}$ are unimodal maps with non-flat critical point. Equilibrium states (in particular of the potential $\varphi_t := -t \log |Df|$) have been studied in this case by various authors [HK2, BK, KN, L, St.P], using transfer operators. The transfer operator, in combination with Markov extensions, proved a powerful tool for so-called Collet-Eckmann unimodal maps (see (3) below) for Keller and Nowicki [KN], who showed that an appropriately weighted version of the transfer operator is quasi-compact. To our knowledge, however, these methods cannot be applied to non-Collet-Eckmann maps.

A less direct approach was taken by Pesin and Senti, results which were announced in [PSe2], with details given in [PSe1]: they used an inducing scheme (X, F, τ) (where τ is the inducing time), a hyperbolic expanding full branched map, albeit with infinitely many branches, to find a unique equilibrium state μ_{Φ_t} for the lifted potential Φ_t. This equilibrium state is then projected to the interval to give a measure μ_{φ_t}, a candidate equilibrium state for the system (I, f, φ_t). It is proved that in the case where f is a unimodal map satisfying the strong exponential growth along critical orbits given in [Se], μ_{φ_t} is a true equilibrium state for the whole system. The down-side for the more general case is that μ_{φ_t} is only an 'equilibrium state' within the class of measures that are *compatible* to the inducing scheme, i.e., the induced map $F = f^\tau$ is defined for all iterates μ-a.e. on X and the inducing time τ is μ_F-integrable (here μ_F is the 'lift' of μ, see below). A priori, the 'equilibrium states' obtained in this way may not be true equilibrium states for the whole system, and different inducing schemes may lead to different measures μ_{φ_t}. Indeed, there exist measures with good properties which lift to some inducing schemes, but not to others: for example if X is small then the set of points which never enter X under iteration by f can support measures of positive entropy. Furthermore, inducing schemes are not always readily available in general.

In this paper we show how to create 'natural' inducing schemes and how to compare measures which 'lift to' different schemes.

Our results are the first to deal with equilibrium states for the potential $\varphi_t : x \mapsto -t \log |Df(x)|$ when f is not Collet-Eckmann. (We emphasise that the corresponding theory in [PSe1] considers a particular set of maps Collet-Eckmann maps close to the Chebychev map.) We also prove results on the analyticity of $t \mapsto P(\varphi_t)$.

The Lyapunov exponent of a measure μ is defined as $\lambda(\mu) := \int_I \log|Df|\, d\mu$. Let $\mathcal{M}_{erg}$ be the set of all ergodic f-invariant probability measures, and

$$\mathcal{M}_+ = \{\mu \in \mathcal{M}_{erg} : \lambda(\mu) > 0,\ \mathrm{supp}(\mu) \not\subset \mathrm{orb}(\mathrm{Crit})\}\,.$$

Measures μ with $\mathrm{supp}(\mu) \subset \mathrm{orb}(\mathrm{Crit})$ are atomic. Atomic measures in $\mathcal{M}_{erg}$ must be supported on periodic cycles. So if $\mathrm{supp}(\mu) \subset \mathrm{orb}(\mathrm{Crit})$ and $\lambda(\mu) > 0$, μ must be supported on a hyperbolic repelling periodic cycle, and thus the corresponding critical point must be preperiodic. (Note that for $t \leqslant 0$ such a situation can produce non-uniqueness of equilibrium states, see [MSm1] and Section 7.)

Theorem 1. *Let $f \in \mathcal{H}$ be transitive with negative Schwarzian derivative and let $\varphi_t := -t\log|Df|$ for $t \in \mathbb{R}$. Suppose that for some $t_0 \in (0,1)$, $C > 0$ and $\beta > \ell_{max}(1 + \frac{1}{t_0}) - 1$,*

$$|Df^n(f(c))| \geqslant Cn^\beta \quad \text{for all } c \in \mathrm{Crit} \text{ and } n \geqslant 1. \tag{1}$$

Then there exists $t_1 \in (t_0, 1)$ such that the following hold:

- *for every $t \in [t_1, 1]$, (I, f, φ_t) has an equilibrium state $\mu_{\varphi_t} \in \mathcal{M}_+$;*
- *if $t_1 < t < 1$, then μ_{φ_t} is the unique equilibrium state in $\mathcal{M}_{erg}$ and a compatible inducing scheme with respect to which μ_{φ_t} has exponential tails;*
- *if $t = 1$, then there may be other equilibrium states in $\mathcal{M}_{erg} \setminus \mathcal{M}_+$. However, for $\mu_{\varphi_1} \in \mathcal{M}_+$ there is a compatible inducing scheme with respect to which μ_{φ_1} has polynomial tails;*
- *the map $t \mapsto P(\varphi_t)$ is analytic on $(t_1, 1)$.*

We refer to this situation as the *summable case*. Note that for $t = 1$ the measure $\mu_{\varphi_1} \in \mathcal{M}_+$ is an absolutely continuous invariant measure (acip). Therefore this result improves on the polynomial case of [BLS, Proposition 4.1], since in that theorem the polynomial decay of the tails was given under the above conditions, but also assuming that the critical points must all have the same order. Results of [BRSS] enable us to drop this assumption. As was shown in [BLS], this tail decay rate implies that the decay of correlations is at least polynomial.

As in the theorem, for $t = 1$ equilibrium states with zero Lyapunov exponent are possible, see Section 7 for details. Let us explain why for $t < 1$, equilibrium states must have $\lambda(\mu) > 0$. The pressure function $t \mapsto P(\varphi_t)$ is a continuous decreasing function. As in [BRSS], condition (1) implies the existence of an acip μ_1 with $\lambda(\mu_1) > 0$, which is also a equilibrium state for the potential $\varphi_1 = -\log|Df|$. It follows that

$$P(\varphi_t) \geqslant (1-t)\lambda(\mu_1) \qquad \text{for all } t \in \mathbb{R}, \tag{2}$$

so if $t < 1$ we have $P(\varphi_t) > 0$. By [Pr], we have $\lambda(\mu) \geqslant 0$ for any invariant measure, so Ruelle's inequality [Ru1] implies that $h_\mu(f) \leqslant \lambda(\mu)$. Thus (for $t < 1$) equilibrium states have positive Lyapunov exponent because $\lambda(\mu) = 0$ implies $P(\varphi_t) = 0$.

Notice that for $t \leqslant 0$, the potential $-t\log|Df|$ is upper semicontinuous, and the entropy function $\mu \mapsto h_\mu(f)$ is upper semicontinuous, as explained in [K2]. This guarantees the existence of equilibrium states for (I, f) when $t \leqslant 0$, regardless of whether (1) holds or not.

A stronger condition than (1) is the *Collet-Eckmann condition* which states that there exist $C, \alpha > 0$ such that

$$|Df^n(f(c))| \geqslant Ce^{\alpha n} \text{ for all } c \in \text{Crit and } n \in \mathbb{N}. \tag{3}$$

This condition implies that $\lambda(\mu) > 0$ for every $\mu \in \mathcal{M}_{erg}$, see e.g. [NS] (and [BS] for the proof in the multimodal case). In the unimodal case, the difference between Collet-Eckmann and non-Collet-Eckmann maps can be seen from the behaviour of the pressure function at $t = 1$, as follows from [NS]. Indeed, if (1) holds but not (3), then there are periodic orbits with Lyapunov exponents arbitrarily close to 0, and hence $P(\varphi_t) = 0$ for $t \geqslant 1$. This is regardless of the existence of equilibrium states, which, for $t > 1$, can only be measures for which $\lambda(\mu) = h_\mu(f) = 0$. This means that the function $t \mapsto P(\varphi_t)$ is not differentiable at $t = 1$: we say that there is a *phase transition* at 1. See Section 7 for more details on the phase transition, and on maps without equilibrium states.

For unimodal Collet-Eckmann maps, the map $t \mapsto P(\varphi_t)$ is analytic in a neighbourhood of 1, as was shown in [BK]. The following theorem (the proof of which introduces many of the ideas used for Theorem 1) generalises this result to all $f \in \mathcal{H}$ satisfying (3), and gives results on equilibrium states also.

Theorem 2. *Suppose $f \in \mathcal{H}$ is transitive with negative Schwarzian derivative and $\varphi_t = -t \log |Df|$. If f is Collet-Eckmann, then there exist $t_1 < 1 < t_2$ such that f has a unique equilibrium state μ_{φ_t} for $t \in (t_1, t_2)$. Moreover, $\mu_{\varphi_t} \in \mathcal{M}_+$, there is a compatible inducing scheme with respect to which μ_{φ_t} has exponential tails, and the map $t \mapsto P(\varphi_t)$ is analytic in (t_1, t_2).*

In fact, the techniques used to prove this theorem also give analyticity of the pressure for the special Collet-Eckmann maps considered in [PSe1] for all t in a neighbourhood of $[0, 1]$.

Lifting measures. Our main theorems deal with equilibrium states in $\mathcal{M}_+$. Although measures in $\mathcal{M}_+$ may not always be compatible to a specific inducing scheme, they are all compatible to some inducing scheme. Given an inducing scheme (X, F, τ), we say that a measure μ_F is a *lift* of μ if for all μ-measurable subsets $A \subset I$,

$$\mu(A) = \frac{1}{\int_X \tau \, d\mu_F} \sum_i \sum_{k=0}^{\tau_i - 1} \mu_F(X_i \cap f^{-k}(A)). \tag{4}$$

Conversely, given a measure μ_F for (X, F), we say that μ_F *projects* to μ if (4) holds.

Let $X^\infty = \cap_n F^{-n}(\cup_i X_i)$ be the set of points on which all iterates of F are defined. The following theorem gives us a method for finding inducing schemes, which are naturally related to measures of positive Lyapunov exponent.

Theorem 3. *If $\mu \in \mathcal{M}_+$, then there is an inducing scheme (X, F, τ) and a measure μ_F on X such that $\int_X \tau \, d\mu_F < \infty$. Here μ_F is the lifted measure of μ (i.e., μ and μ_F are related by (4)). Moreover, if Ω is the transitive component supporting μ then $\overline{X^\infty} = X \cap \Omega$.*

Conversely, if (X, F, τ) is an inducing scheme and μ_F an ergodic F-invariant measure such that $\int_X \tau d\mu_F < \infty$, then μ_F projects to a measure $\mu \in \mathcal{M}_{erg}$ with positive Lyapunov exponent.

We would like to highlight another important set of results in this paper, which will be explained more fully later: We will also show that all 'relevant measures' in this paper lift to a fixed inducing scheme, see Proposition 2 and Lemmas 8 and 10.

The potential φ_t (or $-t\log|Jf|$ in a wider setting, where Jf is the Jacobian of the map) has geometric importance if t is the dimension of the phase space, because then the equilibrium state can often be shown to be absolutely continuous with respect to t-dimensional Hausdorff measure. One can also consider other potentials: e.g. the seminal paper by Bowen [Bo] applies to the class of Hölder potentials. In the setting of interval maps, interesting results and examples were given by Hofbauer and Keller [HK2] for potentials with bounded variation. Our methods extend to such potentials as well. We develop this theory in [BT2].

The paper is organised as follows. Section 2 gives preliminaries on (Gurevich) pressure, recurrence, and gives an important result on symbolic systems, due to Sarig. Also we review basic results for interval maps. Section 3 explains how to find inducing schemes using the Hofbauer tower, which have the important property of being first return map on this tower, even if the inducing scheme is not the first return on the original system (I, f). Theorem 3 is proved here as well. In Section 4 we prove Proposition 1, which gives the basic framework of the existence and uniqueness proofs. Section 5 is devoted to the main part of the proofs of Theorems 1 and 2 (using estimates from [BLS]). In Section 6, we show that most equilibrium states in this paper can be obtained from a Young tower with exponential tails (see [Y] for definitions), and discuss several consequences of this remarkable fact, including the concluding part of Theorems 1 and 2: the analyticity of the pressure function. Finally in Section 7, we discuss the hypotheses of our main theorems and give counter-examples that show that these hypotheses cannot be easily relaxed.

Acknowledgements: We would like to thank Ian Melbourne, Mariusz Urbański, Peter Raith and Benoît Saussol for fruitful discussions and comments on (earlier) versions of this paper. We are especially grateful to Neil Dobbs whose remarks have led to substantial clarification and strengthening of parts of this paper.

2. Preliminaries

2.1. **Measures and Pressure.** If (X, T) is a dynamical system with potential $\Phi : X \to \mathbb{R}$, then the measure m is Φ-*conformal* if

$$m(T(A)) = \int_A e^{-\Phi(x)}\, dm(x)$$

whenever $T : A \to T(A)$ is one-to-one. In other words, $dm \circ T(x) = e^{-\Phi(x)} dm(x)$. We define the transfer operator for the potential Φ as

$$\mathcal{L}_\Phi g(y) := \sum_{T(y)=x} e^{\Phi(y)} g(y).$$

We want to show that whatever inducing scheme we start with, the invariant measure we get on I is unique. One of the key tools is the following theorem which is the main result of [Sa3]. Assume that $\mathcal{S}_1 = \{X_i\}$ is a Markov partition of X such that $T : X_i \to X$ is injective for each $X_i \in \mathcal{S}_1$. We say that (X, T) has the *big images and preimages (BIP)* property if, there exist $X_1, \dots, X_N \in \mathcal{S}_1$ such that for every $X_k \in \mathcal{S}_1$ there are $i, j \in \{1, \dots, N\}$ and $x \in X_i$ such that $T(x) \in X_k$ and $T^2(x) \in X_j$.

Suppose that (X, T) is topologically mixing. For every $X_i \in \mathcal{S}_1$ and $n \geqslant 1$ let

$$Z_n(\Phi, X_i) := \sum_{T^n x = x} e^{\Phi_n(x)} 1_{X_i}(x),$$

where $\Phi_n(x) = \sum_{j=0}^{n-1} \Phi \circ T^j(x)$. Let

$$Z_n^*(\Phi, X_i) := \sum_{\substack{T^n x = x, \\ T^k x \notin X_i \text{ for } 0<k<n}} e^{\Phi_n(x)} 1_{X_i}(x).$$

We define the *Gurevich pressure* of Φ as

$$P_G(\Phi) := \limsup_{n\to\infty} \frac{1}{n} \log Z_n(\Phi, X_i). \tag{5}$$

This limit exists, is independent of the choice of X_i and it is $> -\infty$, see [Sa1]. To simplify the notation, we will often suppress the dependence of $Z_n(\Phi, X_i)$ and $Z_n^*(\Phi, X_i)$ on X_i. Furthermore, if $\|\mathcal{L}_\Phi 1\|_\infty < \infty$ then $P_G(\Phi) < \infty$, see Proposition 1 of [Sa1].

The potential Φ is said to be *recurrent* if

$$\sum_n \lambda^{-n} Z_n(\Phi) = \infty \text{ for } \lambda = \exp P_G(\Phi). \tag{6}$$

Moreover, Φ is called *positive recurrent* if it is recurrent and $\sum_n n\lambda^{-n} Z_n^*(\Phi) = \infty$. We define the *n-th variation* of Φ as

$$V_n(\Phi) := \sup_{\mathbf{C}_n \in \mathcal{S}_n} \sup_{x,y \in \mathcal{S}_n} |\Phi(x) - \Phi(y)|, \tag{7}$$

where $\mathcal{S}_n = \bigvee_{j=0}^{n-1} T^{-j}(\mathcal{S}_1)$ is the n-joint of the Markov partition $\mathcal{S}_1$.

Theorem 4 ([Sa3])**.** *If (X, T) is topologically mixing and $\sum_{n\geqslant 1} V_n(\Phi) < \infty$, then Φ has an invariant Gibbs measure if and only if A has the BIP property and $P_G(\Phi) < \infty$. Moreover the Gibbs measure μ_Φ has the following properties*

(a) *If $h_{\mu_\Phi}(T) < \infty$ or $-\int \Phi d\mu_\Phi < \infty$ then μ_Φ is the unique equilibrium state (in particular, $P(\Phi) = h_{\mu_\Phi}(T) + \int_X \Phi \, d\mu_\Phi$);*
(b) *If $\|\mathcal{L}_\Phi 1\|_\infty < \infty$ then the Variational Principle holds, i.e., $P_G(\Phi) = P(\Phi)$ $(= h_{\mu_\Phi}(T) + \int_X \Phi \, d\mu_\Phi)$;*
(c) *μ_Φ is finite and $\mu_\Phi = \rho_\Phi \, dm_\Phi$ where $\mathcal{L}_\Phi \rho_\Phi = \lambda \rho_\Phi$ and $\mathcal{L}_\Phi^* m_\Phi = \lambda m_\Phi$ for $\lambda = e^{P_G(\Phi)}$, i.e., $m_\Phi(TA) = \int_A e^{\Phi - \log\lambda} \, dm_\Phi$;*

(d) *This* ρ_Φ *is unique and* m_Φ *is the unique* $(\Phi - \log \lambda)$*-conformal probability measure.*

Note that because μ_Φ is a Gibbs measure, $\mu_\Phi(\mathbf{C}_n) > 0$ for every cylinder set $\mathbf{C}_n \in \mathcal{S}_n$, $n \in \mathbb{N}$.

In the paper of Mauldin & Urbański [MU] several similar results can be found, although they use a different approach to pressure, taking the supremum of Φ_n on cylinder sets rather than the value of Φ_n at periodic points.

2.2. **Interval Maps.** An interval map (I, f) is called *piecewise monotone* if there is a finite partition $\mathcal{P}_1$ into maximal intervals on which f is diffeomorphic. We call this partition the *branch partition*. We will assume that f is C^2; negative Schwarzian derivative in this C^2 context means that $1/\sqrt{|Df|}$ is a convex function on each $\mathbf{C} \in \mathcal{P}_1$.

Remark 1. *The negative Schwarzian derivative condition allows us to use the Koebe lemma for distortion control of the branches of the induced maps we obtain later. However if $f \in \mathcal{H}$ is C^3 and there are no neutral periodic cycles, then it is unnecessary to assume negative Schwarzian derivative. This was proved in the unimodal setting by Kozlovski* [Ko]*, and later for $f \in C^{2+\eta}$ in* [T]*. In the multimodal setting for $f \in C^3$ this was proved by van Strien and Vargas* [SV]*.*

Let $\mathcal{P}_n = \bigvee_{k=0}^{n-1} f^{-k}\mathcal{P}_1$. Elements $\mathbf{C}_n \in \mathcal{P}_n$ are called *n-cylinders*. Similarly to (7), the *n-th variation* of a potential $\varphi : I \to \mathbb{R}$ is defined as

$$V_n(\varphi) = \sup_{\mathbf{C}_n \in \mathcal{P}_n} \sup_{x,y \in \mathbf{C}_n} |\varphi(x) - \varphi(y)|.$$

The *non-wandering set* Ω of f is the set of points x having arbitrarily small neighbourhoods U such that $f^n(U) \cap U \neq \emptyset$ for some $n \geqslant 1$. Piecewise monotone C^2 maps have non-wandering sets that split into a finite or countable number of *transitive* components Ω_k such that each Ω_k contains a dense orbit, see [HR] and references therein. A transitive component is one of the following:

(Ω1) A finite union of intervals, cyclically permuted by f. This is the most interesting case, and Lemma 1(a) in Section 3 gives its description on the Hofbauer tower.

(Ω2) A Cantor set if f is *infinitely renormalisable*, i.e, there is an infinite sequence of periodic intervals J_n of increasing periods, and $\Omega = \cap_n \mathrm{orb}(J_n)$. Measures on such components have $\lambda(\mu) = 0$, see [MSt] and [SV, Theorem D] for the multimodal case. For maps that are only piecewise C^2, this is no longer true, see Section 7.

(Ω3) If f is (finitely) renormalisable, say it has a periodic interval $J \neq I$, then the set of points that avoid $\mathrm{orb}(J)$ contains a transitive component as well. This is usually a Cantor set, but it could be a finite set (e.g. if f is the

Feigenbaum map). For infinitely renormalisable maps, there are countably many transitive components of this type. Lemma 1(b) in Section 3 gives its description on the Hofbauer tower.

We will state our results for transitive interval maps, but they can be applied equally well to (Ω_k, f) for any component Ω_k of the non-wandering set. In all our main theorems we assume that (Ω, f) is *topological mixing* (i.e., every iterate of f is topologically transitive). This can be achieved by taking a transitive component of an appropriate iterate of f.

We say that (X, F, τ) is an *inducing scheme* over (I, f) if

- X is a union of intervals containing a (countable) collection of disjoint intervals X_i such that F maps each X_i diffeomorphically onto X, with bounded distortion.
- $F|_{X_i} = f^{\tau_i}$ for some $\tau_i \in \mathbb{N} := \{1, 2, 3 \dots\}$.

The function $\tau : \cup_i X_i \to \mathbb{N}$ defined by $\tau(x) = \tau_i$ if $x \in X_i$ is called the *inducing time*. It may happen that $\tau(x)$ is the first return time of x to X, but that is certainly not the general case. For ease of notation, we will often let $(X, F, \tau) = (X, F)$.

Recall that $X^\infty = \cap_n F^{-n}(\cup_i X_i)$ is the set of points on which all iterates of F are defined. We call a measure μ *compatible* to the inducing scheme if

- $\mu(X) > 0$ and $\mu(X \setminus X^\infty) = 0$, and
- there exists a measure μ_F which projects to μ by (4), and in particular $\int_X \tau \, d\mu_F < \infty$.

Remark 2. (a) *If $\mu \in \mathcal{M}_+$, applying Theorem 3 gives us an inducing scheme (X, F) and a measure μ_F satisfying the above conditions.*

(b) $\overline{X^\infty} = X$ *implies that given a measure μ_F obtained from Theorem 4, the measure μ, the projection of μ_F, has $\mu(U) > 0$ for any open set in $\cup_n f^n(X)$.*

(c) *If (X, F, τ) comes from Theorem 3, then μ is compatible to it if and only if $\mu(X^\infty) > 0$; for more general inducing schemes, this equivalence is false.*

(d) *Note that $\int \tau \, d\mu < \infty$ does not always imply that $\int \tau \, d\mu_F < \infty$, see* [Z].

The inducing scheme (X, F) will perform the role of (X, T) of the previous section, with $\mathcal{S}_1 = \{X_i\}$. Since F maps X_i onto X, the BIP property is automatically satisfied provided F is transitive (if not, we can always select a transitive component). Let us denote the collection of n-cylinders of the inducing scheme by $\mathcal{S}_n$. A priori, $\mathcal{S}_n$ is not connected to $\cup_{m \geqslant 0} \mathcal{P}_m$, i.e., the cylinder sets of the branch partition $\mathcal{P}_1$. In this paper, however, we will always take X to be a subset of $\cup_k \mathcal{P}_k$, and in that case the $\cup_{n \geqslant 1} \mathcal{S}_n \subset \cup_{k \geqslant 1} \mathcal{P}_k$.

Given a potential $\varphi : I \to \mathbb{R}$, let the *lifted potential* Φ be defined by $\Phi(y) = \sum_{j=0}^{\tau_i - 1} \varphi \circ f^j(y)$ for $y \in X_i$. We say that Φ has *summable variations* if $\sum_{n \geqslant 1} V_n(\Phi) < \infty$, and that Φ is *weakly Hölder continuous* if there exist $C_\Phi > 0$ and $0 < \lambda_\Phi < 0$ such that $V_n(\Phi) \leqslant C_\Phi \lambda_\Phi^n$ for all $n \geqslant 1$. Clearly if Φ is weakly Hölder continuous then Φ has summable variations.

We use summability of variations to control distortion of $\Phi_n(x) = \Phi(x) + \cdots + \Phi \circ F^{n-1}(x)$, but for the potential $\varphi_t = -t\log|Df|$, we can also use the Koebe Lemma provided f has negative Schwarzian derivative: If $X' \supset X$ such that X' is a *δ-scaled neighbourhood* of X, i.e., both components of $X' \setminus X$ have length $\geqslant \delta|X|$, and $f^k : X_i \to X$ extends diffeomorphically to $f^k : X_i' \to X'$, then

$$\frac{|Df^k(y)|}{|Df^k(x)|} < \frac{1+2\delta}{\delta^2} + 1$$

for all $x, y \in X_i$.

In this paper we say $A_n \asymp B_n$ if $\lim_{n\to\infty} \frac{A_n}{B_n} = 1$. We will also say that $A \asymp_{dis} B$ if A is equal to B up to some distortion constant.

3. Finding Inducing Schemes

In this section we will prove Theorem 3. The idea relies on the construction of the *canonical Markov extension* $(\hat{I}, \hat{f})$ of the interval map. A measure $\mu \in \mathcal{M}_+$ can be lifted to $(\hat{I}, \hat{f})$, see [K1], and in this space a first return map to a specific subset $\hat{X} \subset \hat{I}$ gives rise to the inducing scheme.

The canonical Markov extension (commonly called *Hofbauer tower*), was introduced by Hofbauer and Keller, see e.g. [H, K1]; it is a disjoint union of subintervals $D = f^n(\mathbf{C}_n)$, $\mathbf{C}_n \in \mathcal{P}_n$, called *domains*, where $\mathcal{P}_1$ is the branch partition. Let $\mathcal{D}$ be the collection of all such domains. For completeness, let $\mathcal{P}_0$ denote the partition of I consisting of the single set I, and call $D_0 = f^0(I)$ the *base* of the Hofbauer tower. Then

$$\hat{I} = \sqcup_{n\geqslant 0} \sqcup_{\mathbf{C}_n \in \mathcal{P}_n} f^n(\mathbf{C}_n)/\sim,$$

where $f^n(\mathbf{C}_n) \sim f^m(\mathbf{C}_m)$ if they represent the same interval. Let $\pi : \hat{I} \to I$ be the inclusion map. Points $\hat{x} \in \hat{I}$ can be written as (x, D) if D is the domain that $\hat{x}$ belongs to and $x = \pi(\hat{x})$. The map $\hat{f} : \hat{I} \to \hat{I}$ is defined as

$$\hat{f}(\hat{x}) = \hat{f}(x, D) = (f(x), D')$$

if there are cylinder sets $\mathbf{C}_n \supset \mathbf{C}_{n+1}$ such that $x \in f^n(\mathbf{C}_{n+1}) \subset f^n(\mathbf{C}_n) = D$ and $D' = f^{n+1}(\mathbf{C}_{n+1})$. In this case, we write $D \to D'$, giving $(\mathcal{D}, \to)$ the structure of a directed graph. It is easy to check that there is a one-to-one correspondence between cylinder sets $\mathbf{C}_n \in \mathcal{P}_n$ and n-paths $D_0 \to \cdots \to D_n$ starting at the base of the Hofbauer tower. For each $R \in \mathbb{N}$, let $\hat{I}_R$ be the compact part of the Hofbauer tower defined by

$$\hat{I}_R = \bigcup \{D \in \mathcal{D} : \text{ there exists a path } D_0 \to \cdots \to D \text{ of length } r \leqslant R\}$$

A subgraph $(\mathcal{E}, \to)$ is called *closed* if $D \in \mathcal{E}$ and $D \to D'$ implies that $D' \in \mathcal{E}$. It is *primitive* if for every pair $D, D' \in \mathcal{E}$, there is a path from D to D' within $\mathcal{E}$. Clearly any two distinct maximal primitive subgraphs are disjoint.

Lemma 1. *Let $f : I \to I$ be a multimodal map and Ω is a transitive component.*
(a) If Ω consists of a finite union of intervals, then there is a closed primitive subgraph $(\mathcal{E}, \to)$ of $(\mathcal{D}, \to)$ containing a dense $\hat{f}$-orbit and such that $\Omega = \pi(\cup_{D\in\mathcal{E}} D)$.

(b) If Ω is a Cantor (or finite) set avoiding a periodic interval of J, then there is a (non-closed) primitive subgraph $(\mathcal{E}, \to)$ of $(\mathcal{D}, \to)$ such that $\Omega \subset \pi(\cup_{D \in \mathcal{E}} D)$, and there is a dense $\hat{f}$-orbit in $(\cup_{D\in\mathcal{E}} D) \cap \pi^{-1}(\Omega)$.

The arguments for this lemma are implicit in [H, HR] combined. We will give a self-contained proof in the appendix. Notice that $(\hat{I}, \hat{f})$ is a Markov map in the sense that the image of any domain D is the union of domains of $\hat{I}$. Obviously, $\pi \circ \hat{f} = f \circ \pi$.

Recall that $D_0 = I = f^0(\mathbf{C}_0)$ is the base of the Hofbauer tower. Let $i : I \to D_0$ be the trivial bijection map (inclusion) such that $i^{-1} = \pi|_{D_0}$. Given a measure $\mu \in \mathcal{M}_{erg}$, let $\hat{\mu}_0 = \mu \circ i^{-1}$, and

$$\hat{\mu}_n := \frac{1}{n} \sum_{k=0}^{n-1} \hat{\mu}_0 \circ \hat{f}^{-k}. \tag{8}$$

We say that μ is *liftable* to $(\hat{I}, \hat{f})$ if there exists a weak accumulation point $\hat{\mu}$ of the sequence $\{\hat{\mu}_n\}_n$ with $\hat{\mu} \not\equiv 0$.

Remark 3. *If μ is liftable and ergodic, then $\hat{\mu}$ is an ergodic $\hat{f}$-invariant probability measure on $\hat{I}$, see* [K1]

Proof of Theorem 3. First assume that $\mu \in \mathcal{M}_+$. Keller [K1] showed that if μ is not atomic then it is liftable, $\hat{\mu}(\hat{I}) = \mu(I) = 1$ and $\hat{\mu} \circ \pi^{-1} = \mu$. If $\mu \in \mathcal{M}_+$ is atomic, it must be supported on a hyperbolic repelling periodic cycle. It is easy to show that such measures are liftable. In both cases, [K1] shows that $\hat{\mu}$ is also ergodic.

Now take some domain D and cylinder set $\mathbf{C}_n \in \mathcal{P}_n$ such that $\pi(D)$ compactly contains $\mathbf{C}_n$ and $\hat{\mu}(\hat{X}) > 0$ for $\hat{X} := \pi^{-1}(\mathbf{C}_n) \cap D$. Let $\hat{F} : \hat{X} \to \hat{X}$ be the first return map; let $\hat{\tau}(x) \in \mathbb{N}$ be such that $\hat{F}(x) = \hat{f}^{\hat{\tau}(\hat{x})}(\hat{x})$ for each $\hat{x} \in \hat{X}$ on which $\hat{F}$ is defined. By the Markov property of $\hat{f}$, $\hat{x}$ has a neighbourhood U such that $\hat{f}^{\hat{\tau}(\hat{x})}$ maps U monotonically onto D. Therefore there is a neighbourhood $V \subset U$ such that $\hat{f}^{\hat{\tau}(\hat{x})}$ maps V monotonically onto $\hat{X}$. Since $\pi(\hat{X}) = \mathbf{C}_n$ is a cylinder set, $\text{orb}(\partial\hat{X}) \cap \hat{X} = \emptyset$. It follows that $\hat{\tau}(\hat{y}) = \hat{\tau}(\hat{x})$ for all $\hat{y} \in V$.

Let Ω be the transitive component supporting μ. If Ω is an interval as in case (Ω1), then we take D inside the closed transitive subgraph of $(\mathcal{D}, \to)$ as guaranteed by Lemma 1(a). Take any open interval $U \subset X$. Since $\mathcal{P}_1$ generates the Borel σ-algebra there is an n-cylinder $\mathbf{C}_n \subset U$; we let $\hat{\mathbf{C}}_n = \pi^{-1}(\mathbf{C}_n) \cap D$. It follows that $\hat{f}^n(\hat{\mathbf{C}}_n) = D'$ for some domain D' in the same transitive component of the Hofbauer tower as D. Hence there is an m-path $D' \to \cdots \to D$ and a subcylinder $\hat{\mathbf{C}}_{n+m} \subset \hat{\mathbf{C}}_n$ such that $\hat{f}^{n+m}(\hat{\mathbf{C}}_{n+m}) = D$. Therefore $\pi(\hat{\mathbf{C}}_{n+m}) \subset U$ contains a domain X_i. It follows that $\cup_i X_i$ is dense in X. Repeating the argument for $U \subset X_i$ we find that $F^{-1}(\cup_i X_i)$ is dense in X, and by induction, X^∞ is dense in X as well. (Notice that this construction may produce many branches X_i such that $\mu(X_i) = 0$, but this doesn't affect the result.)

If Ω is as in case (Ω2) then $\mathcal{M}_+ = \emptyset$ so there is nothing to show. This is proved for the unimodal case in [MSt]; the multimodal case is similar, the required 'real bounds' follow from [SV]. If Ω is Cantor (or finite) set of points avoiding a periodic interval of f as in case (Ω3), then Lemma 1(b) still provides us with a primitive subgraph, and the same argument as above shows that X^∞ is dense in $X \cap \Omega$.

Now the inducing scheme (X, F, τ) is defined by $X = \pi(\hat{X})$, $F = \pi \circ \hat{F} \circ \pi^{-1}|_{\hat{X}}$ and $\tau(x) = \hat{\tau}(\pi^{-1}(x) \cap \hat{X})$. Because $\mu = \hat{\mu} \circ \pi^{-1}$, $\mu(X) \geqslant \hat{\mu}(\hat{X}) > 0$.

Let $\hat{\mu}_{\hat{X}} := \frac{1}{\hat{\mu}(\hat{X})}\hat{\mu}|_{\hat{X}}$ be the conditional measure on $\hat{X}$. The measure $\mu_F := \hat{\mu}_{\hat{X}} \circ \pi^{-1}|_{\hat{X}}$ is clearly F-invariant, and by Kac's Lemma,

$$\int_X \tau \, d\mu_F = \int_{\hat{X}} \hat{\tau} \, d\hat{\mu}_{\hat{X}} = \frac{1}{\hat{\mu}(\hat{X})} < \infty.$$

Finally, by the Poincaré Recurrence Theorem, $\hat{\mu}_{\hat{X}}$-a.e. point $\hat{x} \in \hat{X}$ returns infinitely often to $\hat{X}$, and because $\mu_F \ll \mu$ we also get $\mu(X^\infty) = \mu(X)$ by ergodicity.

Now for the other direction, notice that by assumption, each branch of any iterate F^n of the induced map has negative Schwarzian derivative. Therefore distortion is bounded uniformly over n and the branches of F^n. Hence, by taking an iterate of the induced map F if necessary, we can assume that F^n is uniformly expanding. It follows by F-invariance of μ_F that

$$\begin{aligned} 0 &< \frac{1}{n}\int_{X^\infty} \log|DF^n| \, d\mu_F \\ &= \int_{X^\infty} \frac{1}{n}\sum_{j=0}^{n-1} \log|DF \circ F^j| \, d\mu_F = \int_{X^\infty} \log|DF| \, d\mu_F = \lambda(\mu_F). \end{aligned}$$

Let μ be the projected measure of μ_F; both μ_F and μ are ergodic. Since $\int \tau \, d\mu_F < \infty$, we can take a point $x \in X^\infty$ which is typical for both μ_F and μ. Let $\tau_k = \sum_{j=0}^{k-1} \tau \circ F^j(x)$. Then applying the Ergodic Theorem several times, we get $\lim_{k\to\infty} \frac{\tau_k}{k} = \int \tau d\mu_F < \infty$ and

$$\begin{aligned} \lambda(\mu) &= \int_I \log|Df| \, d\mu = \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \log|Df \circ f^j(x)| \\ &= \lim_{k\to\infty} \frac{1}{\tau_k}\sum_{j=0}^{\tau_k-1} \log|Df \circ f^j(x)| \\ &= \lim_{k\to\infty} \frac{k}{\tau_k}\,\frac{1}{k}\sum_{j=0}^{k-1} \log|DF \circ F^j(x)| = \frac{1}{\int \tau d\mu_F}\lambda(\mu_F) > 0. \end{aligned}$$

This concludes the proof. □

Remark 4. *If* $\lambda(\mu) > 0$ *but* $\mathrm{supp}(\mu) \subset \mathrm{orb}(\mathrm{Crit})$ *and* μ *is the equidistribution on a repelling periodic orbit, say* $\mathrm{supp}(\mu) = \mathrm{orb}(p)$ *where* $f^n(p) = p$, *then we can still find an inducing scheme compatible to* μ. *Let* $X \ni p$ *be an open interval such that the component of* $f^{-n}(X)$ *containing* p *is compactly contained in* X. *Call this*

component X_1. Then (X, F, τ) with $F|_{X_1} = f^{\tau_1}|_{X_1} = f^n|_{X_1}$ is an inducing scheme compatible to μ.

Remark 5. *If $\mu \in \mathcal{M}_+$ then Remark 3 implies that $\hat\mu$ is ergodic. If Ω is as in Lemma 1(a) we also have that $\hat\mu$ is supported on $\mathcal{E}$. That lemma implies that for any $\hat x \in \hat I \setminus \partial\mathcal{D}$ there is $\hat y \in \mathcal{E}$ so that $\pi(\hat x) = \pi(\hat y)$. Thus there exists $n \geqslant 0$ so that $\hat f^n(\hat x) = \hat f^n(\hat y)$. So $\hat\mu(\mathcal{E}) = 1$ follows by ergodicity.*

The induced system used in this proof may be the simplest but not always the most convenient. Let us call an inducing scheme (X, F, τ) a *first extendible return* scheme with respect to a neighbourhood Y of X if for each $x \in X_i$, $\tau(x)$ is the smallest positive iterate such that $f^j(x) \in X$ and there is a neighbourhood $Y_i \supset X_i$ such that f^j maps Y_i monotonically onto Y. If Y is a fixed δ-scaled neighbourhood Y, then the Koebe Lemma can be used to control distortion of branches of (iterates of) F. In this case we say that τ is the *first δ-extendible return time to X.*

Lemma 2. *If $\mu \in \mathcal{M}_+$ then there exists $\delta > 0$ and an interval $X \subset I$ such that μ is compatible to the inducing scheme (X, F, τ) where τ is the first δ-extendible return time. Moreover, if Ω is the transitive component supporting μ then $\overline{X^\infty} = X \cap \Omega$.*

The proof of the first part of this lemma can be found in [B1], but some of the ideas of the proof are particularly useful in this paper so we sketch those parts here.

Proof. As we noted in the proof of Theorem 3, since $\mu \in \mathcal{M}_+$, $\hat\mu(\hat I) > 0$. We choose X and $\delta > 0$ so that the set $\hat X = \sqcup\{D \cap \pi^{-1}(X) : D \in \mathcal{D}, \pi(D) \supset Y\}$, where Y is concentric with X and size $(1+2\delta)|X|$, has $\hat\mu(\hat X) > 0$. Let $r_{\hat X}$ denote the first return map to $\hat X$. In [B1] it is shown that given $x \in X^\infty$, for any $\hat x \in \hat X$ with $\pi(\hat x) = x$, we have $r_{\hat X}(\hat x) = \tau(x)$. As in [B1], this can be used to prove that μ is compatible to (X, F, τ).

The proof that $\overline{X^\infty} = X \cap \Omega$ follows as in the proof of Theorem 3. □

Theorem 3 exploits the fact that measures with positive Lyapunov exponents are liftable; but their lifts do not, in general, give similar mass to the same parts in the Hofbauer tower. The next result shows that measures with entropy uniformly bounded away from 0 lift, and give mass uniformly to specific compact subsets of the Hofbauer tower. The proof is postponed to the appendix.

Lemma 3. *For every $\varepsilon > 0$, there are $R \in \mathbb{N}$ and $\eta > 0$ such that if $\mu \in \mathcal{M}_{erg}$ has entropy $h_\mu(f) \geqslant \varepsilon$, then μ is liftable to the Hofbauer tower and $\hat\mu(\hat I_R) \geqslant \eta$. Furthermore, there is a set $\hat E$, depending only on ε, such that $\hat\mu(\hat E) > \eta/2$ and $\min_{D \in \mathcal{D} \cap \hat I_R} d(\hat E \cap D, \partial D) > 0$.*

One consequence of this lemma is that the choice of δ in Lemma 2 depends only on the entropy of μ.

Notice that by Remark 5, we can suppose that $\hat E \subset \mathcal{E}$. We will use this lemma in connection with Case 4 of Proposition 1 in the next section to carry out the proofs

of Theorems 2 and 1. In principle, these results deal with measures in $\mathcal{M}_+$ that possibly have zero entropy. However, the next lemma shows that our equilibrium states need to have both positive Lyapunov exponent and entropy.

Lemma 4. *Suppose that $f \in \mathcal{H}$ satisfies* (1). *Then there exists $\zeta_1 < 0$ so that for $t \in (\zeta_1, 1)$, there exist $\varepsilon_0, \varepsilon > 0$ so that any measure ν with $h_\nu(f) + \int \varphi_t \, d\nu > P(\varphi_t) - \varepsilon_0$ satisfies $h_\nu(f) \geqslant \varepsilon$. Similarly, if $f \in \mathcal{H}$ satisfies* (3) *then there exist $\zeta_1 < 0 < \zeta_2$ so that for $t \in (\zeta_1, 1+\zeta_2)$, there exist $\varepsilon_0, \varepsilon > 0$ so that any measure ν with $h_\nu(f) + \int \varphi_t \, d\nu > P(\varphi_t) - \varepsilon_0$ satisfies $h_\nu(f) \geqslant \varepsilon$.*

Proof. Any transitive map satisfying (1) has an acip μ with $h_\mu(f) = \lambda(\mu) > 0$. Applying (2) and Ruelle's inequality [Ru1], we obtain that $P(\varphi_t) > 0$ for $t < 1$. We let $\varepsilon_0 = \varepsilon_0(t) := P(\varphi_t)/2$. Therefore, it is easy to see that for all $t \in [0, 1)$ there exists $\varepsilon = \varepsilon(t) > 0$ such that $h_\nu(f) + \int \varphi_t \, d\nu > P(\varphi_t)/2$ implies $h_\nu(f) > \varepsilon$. For the case $t < 0$, let $\zeta_1 := -\frac{h_{top}(f)}{4\sup\{\lambda(\nu) : \nu \in \mathcal{M}_{erg}\}}$. Then $h_\nu(f) + \int \varphi_t \, d\nu > P(\varphi_t)/2$ implies $h_\nu(f) > P(\varphi_t)/2 - t\lambda(\nu)$. Since $P(\varphi_t) > h_{top}(f)$, for $t \in (\zeta_1, 0)$ we obtain $h_\nu(f) > h_{top}(f)/4$.

Next assume that the Collet-Eckmann condition (3) holds. We can choose ζ_1 as above. Define $\underline{\lambda} := \inf\{\lambda(\nu) : \nu \in \mathcal{M}_{erg}\}$, and let $\gamma := \underline{\lambda}/\lambda(\mu) \leqslant 1$. By [BS, Theorem 1.2] we know that $\underline{\lambda} > 0$. Take $\varepsilon = \underline{\lambda}/2$. If ν is any measure with $h_\nu(f) < \varepsilon$ then

$$P(\varphi_t) - \left(h_\nu(f) + \int \varphi_t d\nu\right) \geqslant \left[(1-t) - \left(\frac{1}{2} - t\right)\gamma\right] \lambda(\mu) = \left[1 - \frac{\gamma}{2} + t(\gamma - 1)\right] \lambda(\mu),$$

which is bounded away from 0 for all fixed $1 \leqslant t < \frac{1-\gamma/2}{1-\gamma}$ (or all $t \geqslant 1$ if $\gamma = 1$). Hence, if $h_\nu(f) < \varepsilon$, then the free energy of ν cannot be close to $P(\varphi_t)$. □

We are now able to state the following, which relates to part (c) of Proposition 1.

Corollary 1. *In the setting of Theorems 1 and 2, there exists $\eta' > 0$, a sequence $\{\mu_n\}_n$ such that $h_{\mu_n}(f) + \int \varphi_t \, d\mu_n \to P(\varphi_t)$ and an inducing scheme (X, F) given by Theorem 3 or a first extendible return map (as in Lemma 2) such that $\hat\mu_n(\hat X) > \eta'$ for all n.*

Proof. From the definition of pressure, there exists $\{\mu_n\} \subset \mathcal{M}_{erg}$ so that $h_{\mu_n}(f) + \int \varphi_t \, d\mu_n \to P(\varphi_t)$. By Lemma 4, there exists $\varepsilon > 0$ so that $h_{\mu_n}(f) \geqslant \varepsilon$ for all large n. Let $\hat E = \hat E(\varepsilon)$ as in Lemma 3. Firstly, for the type of inducing scheme given by Theorem 3, there must exist $\eta' > 0$, $D \in \mathcal{D} \cap \hat I_R$, a subset $\hat E' \subset \hat E \cap D$ with $\pi(\hat E') \in \mathcal{P}_n$ and a subsequence $n_k \to \infty$ such that $\mu_{n_k}(\hat E') \geqslant \eta'$. Then we let $\hat E'$ be the inducing domain $\hat X$ in Theorem 3. Lemmas 3 and 4 complete the proof.

For a first extendible inducing scheme as in Lemma 2, the proof follows similarly. The main point is to notice that the set $\hat E$ from Lemma 3 has $\min_{D \in \mathcal{D} \cap \hat I_R} d(\hat E \cap D, \partial D) > 0$. □

4. A Key Result for Existence and Uniqueness

The proof of Theorem 1 is divided into several steps. We use the Hofbauer tower construction given in Section 3 to fix an inducing scheme $F : \bigcup_j X_j \to X$ over $X \in \mathcal{P}_n$. Let Φ be the induced potential.

The following lemma, the ideas for which go back to Abramov [Ab], relates the free energies of the original and the induced system. See [PSe1] for the proof.

Lemma 5. *If μ_F is an ergodic measure on (X, F) with $\int \tau d\mu_F < \infty$, and μ is the projected measure on (X, f), then*

$$h_{\mu_F}(F) = \left(\int_X \tau \; d\mu_F\right) h_\mu(f) \text{ and } \int_X \Phi \; d\mu_F = \left(\int_X \tau \; d\mu_F\right) \int_I \varphi \; d\mu.$$

where Φ is the lifted potential of φ.

It is easy to show that putting $\varphi := \log |Df|$ into the above lemma proves that for any full-branched inducing scheme with ergodic invariant measure μ_F, the measure projects to a measure μ with $\lambda(\mu) > 0$.

Suppose that $\varphi : I \to \mathbb{R}$ is the potential for the original system. We will deal with the shifted potential $\psi_S := \varphi - S$. Given an inducing scheme (X, F) with $F = f^\tau$, let Ψ_S be the induced potential, i.e., $\Psi_S := \Phi - \tau S$. The following lemma resembles the argument of [Sa1, Proposition 10]. An important difference here is that we do not require that the original potential has summable variations.

Lemma 6. *Suppose that $P_G(\Psi_{S^*}) < \infty$ and Φ has summable variations. Then $P_G(\Psi_S)$ is decreasing and continuous in $[S^*, \infty)$.*

Proof. We first recall some facts. By definition, $P_G(\Psi_S) := \lim_{n\to\infty} \frac{1}{n} \log Z_n(\Psi_S, X_i)$ where $Z_n(\Psi_S, X_i) := \sum_{F^n x = x} e^{(\Psi_S)_n(x)} 1_{X_i} = \sum_{F^n x = x} e^{\Phi_n(x) - S\tau^n(x)} 1_{X_i}$. As in [Sa1], topological mixing implies that $P_G(\Psi_S)$ is independent of X_i, and we suppress X_i in the notation accordingly. Clearly, $P_G(\Psi_S)$ is decreasing in S. We also know that since we have summable variations for Φ, i.e., there exists $B < \infty$ such that $\sum_{k=1}^\infty V_n(\Phi) < B$, we have for any S,

$$\log Z_{m_1}(\Psi_S) + \log Z_{m_2}(\Psi_S) \leqslant \log Z_{m_1+m_2}(\Psi_S) + \log B, \tag{9}$$

see the proof of [Sa1, Proposition 1].

Since $P_G(\Psi_S)$ is decreasing in S, it is sufficient to show that for any $S_0 \geqslant S^*$ and any $\varepsilon > 0$, there exists $S > S_0$ such that $P_G(\Psi_S) > P_G(\Psi_{S_0}) - \varepsilon$. Fix $\varepsilon > 0$ and n_0 so large that $\frac{\log B}{n_0} < \frac{\varepsilon}{3}$. By definition of $P_G(\Psi_{S_0})$, for a large enough $n \geqslant n_0$,

$$\frac{1}{n} \log Z_n(\Psi_{S_0}) \geqslant P_G(\Psi_{S_0}) - \frac{\varepsilon}{3}.$$

Since $Z_n(\Psi_S)$ is continuous in S, there exists $S > S_0$ such that

$$\frac{1}{n} \log Z_n(\Psi_S) > P_G(\Psi_{S_0}) - \frac{2}{3}\varepsilon.$$

Then by (9) and writing $m = kn + r$ where $0 \leqslant r \leqslant n-1$,

$$\frac{\log Z_m(\Psi_S)}{m} \geqslant \frac{k\log Z_n(\Psi_S) + \log Z_r(\Psi_S) - (k+1)\log B}{kn+r}$$
$$\overset{m\to\infty}{\longrightarrow} \frac{\log Z_n(\Psi_S)}{n} - \frac{\log B}{n} \geqslant P_G(\Psi_{S_0}) - \varepsilon$$

as required. □

The following result is a key tool in proving Theorems 1 and 2. It gives necessary conditions, comparable to the abstract conditions presented in [PSe1], to push equilibrium states through inducing procedures. Notice that Case 4 is reminiscent of the ideas involved in the Discriminant Theorem, [Sa2, Theorem 2]. However, our approach seems more natural in this context.

Proposition 1. *Suppose that ψ is a potential with $P(\psi) = 0$. Let $\hat{X}$ be the set used in either Theorem 3 or Lemma 2 to construct the corresponding inducing scheme (X, F, τ). Suppose that the lifted potential Ψ has $\|\mathcal{L}_\Psi 1\|_\infty < \infty$ and $\sum_{n\geqslant 1} V_n(\Psi) < \infty$.*

Consider the assumptions:

(a) $\sum_i \tau_i e^{\Psi_i} < \infty$ *for* $\Psi_i = \sup_{x\in X_i} \Psi(x)$*;*
(b) *there exists an equilibrium state $\mu \in \mathcal{M}_+$ compatible to (X, F, τ);*
(c) *there exist a sequence $\{\varepsilon_n\}_n \subset \mathbb{R}^-$ with $\varepsilon_n \to 0$ and measures $\{\mu_n\}_n \subset \mathcal{M}_+$ such that every μ_n is compatible to (X, F, τ), $h_{\mu_n}(f) + \int \psi \, d\mu_n = \varepsilon_n$ and $P_G(\Psi_{\varepsilon_n}) < \infty$ for all n;*
(d) $P_G(\Psi) = 0$.

If any of the following combinations of assumptions holds:

$$\begin{cases} 1. & \textit{(b) and (d)};\\ 2. & \textit{(a) and (d)};\\ 3. & \textit{(a) and (b)};\\ 4. & \textit{(a) and (c)}; \end{cases}$$

then there is a unique equilibrium state μ for (I, f, ψ) among measures $\mu \in \mathcal{M}_+$ with $\hat{\mu}(\hat{X}) > 0$. Moreover, μ is obtained by projecting the equilibrium state μ_Ψ of the inducing scheme and in all cases we have $P_G(\Psi) = 0$.

Remark 6. *As noted in the proof, if μ_Ψ is the equilibrium state for (X, F, Ψ) given by Theorem 4 then the condition $\sum_i \tau_i e^{\Psi_i} < \infty$ implies that $\int_Y \tau d\mu_\Psi < \infty$ by the Gibbs property of μ_Ψ.*

Proof of Proposition 1. As in Section 2, Proposition 1 of [Sa1] implies that $Z_n(\Psi) = O(\|\mathcal{L}_\Psi 1\|_\infty^n)$. Therefore $\|\mathcal{L}_\Psi 1\|_\infty < \infty$ implies $P_G(\Psi) < \infty$. So in any case we can immediately apply Theorem 4 to obtain a measure μ_Ψ, and moreover the Variational Principle holds.

Case 1. (b) and (d) hold: By definition of compatibility, we can lift μ to μ_F where $\int \tau \, d\mu_F < \infty$. By Lemma 5 we have

$$0 = P(\psi) = \left(\int \tau \, d\mu_F\right)\left(h_\mu(f) + \int \psi \, d\mu\right) = h_{\mu_F}(F) + \int \Psi \, d\mu_F.$$

Since we also have $P_G(\Psi) = 0$, the Variational Principle (Theorem 4 (b)) implies that μ_F is an equilibrium state for the inducing scheme. From the uniqueness of the measure given by Theorem 4, we have $\mu_F = \mu_\Psi$. So μ is the same as the projection of μ_Ψ given by Theorem 3, as required. Note that by Lemma 5, $h_{\mu_\Psi}(F) < \infty$ and $-\int \Psi \, d\mu_\Psi < \infty$.

Case 2: (a) and (d) hold: By the Gibbs property of μ_Ψ we have

$$\int \tau \, d\mu_\Psi \asymp_{dis} \sum_i \tau_i e^{\Psi_i - P_G(\Psi)} < \infty.$$

This implies that we can use Theorem 3 to project μ_Ψ to an f-invariant measure $\mu_\psi \in \mathcal{M}_+$. By Lemma 5, $h_{\mu_\Psi}(F) < \infty$ and $-\int \Psi \, d\mu_\Psi < \infty$. So by Theorem 4 part (a), μ_Ψ is an equilibrium, and the Variational Principle (i.e., Theorem 4 part (b)) we have $P_G(\Psi) = P(\Psi) = h_{\mu_\Psi}(F) + \int \Psi \, d\mu_\Psi$.

Now condition (d) gives that $P_G(\Psi) = P(\Psi) = 0$. Thus Lemma 5 implies that $h_{\mu_\psi}(f) + \int \psi \, d\mu_\psi = 0$, so μ_ψ is an equilibrium state. We can then use the argument of Case 1 to show that this is the unique equilibrium state in $\mathcal{M}_+$ with $\hat\mu(\hat X) = (\int \tau \, d\hat\mu)^{-1} > 0$.

Case 3: (a) and (b) hold: We start as in Case 2; condition (a) gives a measure μ_ψ having $h_{\mu_\psi}(f) + \int \psi \, d\mu_\psi \leqslant P(\psi) = 0$. By Lemma 5 and the Variational Principle this implies $P_G(\Psi) \leqslant 0$.

Assumption (b) gives an equilibrium state $\mu \in \mathcal{M}_+$ which can be lifted, using Theorem 3, to μ_F on (X, F, τ). Now since we also have $0 = h_\mu(f) + \int \psi \, d\mu$, Lemma 5 implies that $0 \leqslant \int \tau \, d\mu_F(h_\mu(f) + \int \psi \, d\mu) \leqslant P(\Psi)$ and by the Variational Principle, $0 \leqslant P_G(\Psi)$ as well. Thus we have $P_G(\Psi) = 0$ and we can apply the argument of Case 1.

Case 4: (a) and (c) hold: By the argument of Case 2 we have an equilibrium state μ_ψ. Therefore, if we can show that $P_G(\Psi) = 0$, Case 1 above completes the proof.

The argument for Case 3 showed that $P_G(\Psi) \leqslant 0$. By (c), $h_{\mu_n}(f) + \int (\psi - \varepsilon_n) \, d\mu_n = -\varepsilon_n > 0$. Let $\mu_{n,F}$ be the corresponding lifted measure obtained from Theorem 3. Then by Lemma 5, $0 \leqslant h_{\mu_{n,F}}(F) + \int_X \Psi_{\varepsilon_n} \, d\mu_{n,F} \leqslant P_G(\Psi_{\varepsilon_n})$. Lemma 6 implies that we can take the limit to get $P_G(\Psi) = \lim_{n \to \infty} P_G(\Psi_{\varepsilon_n}) = 0$. □

We next present a technical result, which when applied to the settings of Theorems 1 and 2, shows that any measure with free energy close to our equilibrium states lifts to a single inducing scheme, see Lemma 10.

Lemma 3 says that given $\varepsilon > 0$ there exists $\eta = \eta(\varepsilon)$ and $\hat{E} = \hat{E}(\varepsilon)$, a compact set bounded away from $\partial\mathcal{D}$, so that $h_\mu(f) > \varepsilon$ for $\mu \in \mathcal{M}$ implies $\hat{\mu}(\hat{E}) > \eta$. This implies that for a measure $\mu \in \mathcal{M}_+$, in particular an equilibrium state μ_ψ, we can choose $X^0 \in \mathcal{P}_n$ so that for the set $\hat{X}^0$ as in Theorem 3 (or Lemma 2 if a first extendible return map is preferred) $\hat{\mu}_\psi(\hat{X}^0 \cap \hat{E}) > 0$. Next we add a finite collection of cylinder sets $X^k \in \cup_{j \geqslant n}\mathcal{P}_j$, $k = 1, \dots, N$, so that if we create the sets $\hat{X}^k \subset \pi^{-1}(X^k)$ in the same way (i.e., as in Theorem 3 or as in Lemma 2), then $\hat{E} \subset \left(\cup_{0 \leqslant k \leqslant N}\hat{X}^k\right)$. In this case we say that $\{\hat{X}^k\}_{0 \leqslant k \leqslant N}$ satisfies property $Cover(\varepsilon)$. The next proposition shows that there is a single inducing scheme that is compatible to every measure in $\mathcal{M}_+$ whose free energy is sufficiently close to the pressure.

Proposition 2. *Suppose that $\psi : I \to [-\infty, \infty)$ is a potential with $P(\psi) = 0$ so that $\psi(x) > -\infty$ on $I \setminus \mathrm{Crit}$. Suppose also that there exist $\varepsilon_0, \varepsilon > 0$ such that $h_{\mu'}(f) + \int \psi \, d\mu' > -\varepsilon_0$ implies $h_{\mu'}(f) > \varepsilon$. Let $\{\hat{X}^k\}_{0 \leqslant k \leqslant N}$ satisfy $Cover(\varepsilon)$ where μ_ψ is compatible to (X^0, F_0). Suppose that the induced potentials Ψ^k and inducing times τ^k corresponding to the inducing schemes (X^k, F_k) satisfy:*

(a) $\sum_n V_n(\Psi^k) < \infty$ *for all* $0 \leqslant k \leqslant N$;

(b) $\sum_i \tau_i^k e^{\sup\{\Psi^k(x) : x \in X_i^k\}} < \infty$ *(i.e., condition (a) of Proposition 1 holds for Ψ^k) for all $0 \leqslant k \leqslant N$.*

Then there exists $\theta = \theta(\varepsilon, \{\hat{X}^k\}_{0 \leqslant k \leqslant N}) > 0$ so that $h_\mu(f) + \int \psi \, d\mu > -\theta$ implies $\hat{\mu}(\hat{X}_0) > 0$.

The idea here is that information on the equilibrium state for (X^0, F_0, Ψ^0) allows us to show that measures with enough free energy must cover a large portion of the Hofbauer tower, in particular they are compatible to (X^0, F_0).

Proof. Let $k \in \{1, \dots, N\}$ be arbitrary and assume that $\mu' \in \mathcal{M}_+$ is a measure such that $\hat{\mu}'(\hat{X}^k) > 0$, but with $\hat{\mu}'(\hat{X}^0) = 0$.

Here we will refer to the components of $\pi^{-1}(X_i^k) \cap \hat{X}^k$ as 1-*cylinders of* $(\hat{X}^k, R_{\hat{X}^k})$, the first return map to $\hat{X}^k$.

Claim 1. (i) *There is at least one* 1*-cylinder mapping into $\hat{X}^0$ before returning to $\hat{X}^k$;*

(ii) *There is at least one* 1*-cylinder which does not map to $\hat{X}^0$ before returning to $\hat{X}^k$.*

Moreover, whether (i) or (ii) holds depends only on $\pi(\hat{X}_i^k)$, and not on the domain that $\hat{X}_i^k$ belongs to.

Proof. Property (i) follows by transitivity. (A priori, sets $\hat{X}_i^k$ satisfying (i) may have $\hat{\mu}'(\hat{X}_i^k) = 0$ or not; we will show that$\hat{\mu}'(\hat{X}_i^k) > 0$ for at least one such $\hat{X}_i^k$.)

For property (ii), suppose that for any first return domain $\hat{X}_i^k \subset D \in \mathcal{D}$ there is $0 \leqslant s < r_{\hat{X}^k}(\hat{X}_i^k)$ such that $\hat{f}^s(\hat{X}_i^k) \cap \hat{X}^0 \neq \emptyset$. By the properties of cylinders we must in fact have $\hat{f}^s(\hat{X}_i^k) \subset \hat{X}^0$. This means that $\hat{\mu}'$-a.e. point enters $\hat{X}^0$ with positive frequency. Ergodicity implies that $\hat{\mu}'(\hat{X}^0) > 0$ which is a contradiction. Hence (ii) holds.

Since $\hat{X}^k \in \cup_{j \geqslant n} \mathcal{P}_j$, if (i) holds for some 1-cylinder $\hat{X}_i^k$ of $(\hat{X}^k, R_{\hat{X}^k})$, say, then this whole cylinder maps into $\hat{X}^0$. Moreover, by the proof of Lemma 2, see [B1], if $\hat{y}_1, \hat{y}_2 \in \hat{X}^k$ have $\pi(\hat{y}_1) = \pi(\hat{y}_2)$ and $\hat{f}^k(\hat{y}_1) \in \hat{X}^0$ then $\hat{f}^k(\hat{y}_2) \in \hat{X}^0$. Consequently, for a 1-cylinder X_i^k of (X^k, F_k) either every component of $\pi^{-1}(X_i^k) \cap \hat{X}^k$ has property (i), or every component of $\pi^{-1}(X_i^k) \cap \hat{X}^k$ has property (ii). This concludes the proof of the first claim. □

Since, by the Gibbs property from Theorem 4, μ_Ψ gives all cylinders of (X^0, F_0) positive mass, the same must be true of the $\hat{\mu}_\psi \circ \pi|_{\hat{X}^0}^{-1}$-measure of these cylinders. Thus part (i) of the claim implies that $\hat{\mu}_\psi(\hat{X}^k) > 0$ and hence μ_ψ is compatible to (X^k, F_k). By Case 3 of Proposition 1, this also implies that $P_G(\Psi^k) = 0$.

Let $(X_\flat^k, F_k)$ denote the system minus the cylinders satisfying (i). Let $P_G^\flat(\Psi^k)$ denote the Gurevich pressure of $(X_\flat^k, F_k, \Psi^k)$, computed from $Z_n^\flat(\Psi^k)$, which is defined in the natural way. (Note that one consequence of part (ii) of the claim is that $P_G^\flat(\Psi^k) > -\infty$.)

Claim 2. $P_G^\flat(\Psi^k) < P_G(\Psi^k) = 0$.

Proof. Let $\mathcal{Y}^k$ be the union of 1-cylinders of (X^k, F_k) whose representatives in $\hat{X}^k$ satisfy property (i). We fix a 1-cylinder Y^k so that $Y^k \cap \mathcal{Y}^k = \emptyset$, i.e., its representatives in $\hat{X}^k$ satisfy (ii). In each $\mathbf{C}_j^k \subset Y^k$ there exists a unique periodic point which contributes to $Z_j(\Psi^k, Y^k)$. Thus noting that $m_{\Psi^k}(\mathbf{C}_j^k) = \int_{\mathbf{C}_j^k} e^{-\Psi^k(x)} d\mu_{\Psi^k}$ and using the variation properties of Ψ_j^k, we derive

$$e^{-V_j(\Psi^k)} \sum m_{\Psi^k}(\mathbf{C}_j^k) \leqslant Z_j(\Psi^k, Y^k) \leqslant e^{V_j(\Psi^k)} \sum m_{\Psi^k}(\mathbf{C}_j^k)$$

where the sum is taken over all j-cylinders $\mathbf{C}_j^k$ in Y^k. Similarly

$$e^{-V_j(\Psi^k)} \sum{}^\flat m_{\Psi^k}(\mathbf{C}_j^k) \leqslant Z_j^\flat(\Psi^k, Y^k) \leqslant e^{V_j(\Psi^k)} \sum{}^\flat m_{\Psi^k}(\mathbf{C}_j^k)$$

where the sum $\sum^\flat$ is taken over all j-cylinders $\mathbf{C}_j^k$ in Y^k so that $F_k^s(\mathbf{C}_j^k) \cap \mathcal{Y}^k = \emptyset$ for $0 \leqslant s \leqslant j-1$.

For every $\mathbf{C}_j^k$ in the sum $\sum^\flat m_{\Psi'}(\mathbf{C}_j^k)$ there exist collection of $j+1$-cylinders $\mathbf{C}_{j+1}^k$ so that $F_k^j(\cup \mathbf{C}_{j+1}^k) = \mathcal{Y}^k$. Since m_{Ψ^k} is conformal and Ψ^k has summable variations, we have

$$\frac{m_{\Psi^k}(\cup \mathbf{C}_{j+1}^k)}{m_{\Psi^k}(\mathbf{C}_j^k)} \geqslant \frac{1}{K} \left(\frac{m_{\Psi^k}(\mathcal{Y}^k)}{m_{\Psi^k}(X^k)} \right)$$

where $K = e^{\sum_j V_j(\Psi^k)}$. Hence, since $m_{\Psi^k}(X^k) = 1$,

$$\begin{aligned}\sum\nolimits^{\flat} m_{\Psi^k}(\cup \mathbf{C}^k_{j+1}) &= \sum\nolimits^{\flat}(m_{\Psi^k}(\mathbf{C}^k_j) - m_{\Psi^k}(\cup \mathbf{C}^k_{j+1})) \\ &\leqslant \left(1 - \frac{m_{\Psi^k}(\mathcal{Y}^k)}{K}\right)\sum\nolimits^{\flat} m_{\Psi^k}(\mathbf{C}^k_j).\end{aligned}$$

Letting $\xi := \frac{\mu_{\Psi^k}(\mathcal{Y}^k)}{K}$ we have

$$Z^{\flat}_{j+1}(\Psi^k, Y^k) \leqslant e^{V_{j+1}(\Psi^k)} \sum\nolimits^{\flat} \mu_{\Psi^k}(\mathbf{C}^k_j) \leqslant e^{V_{j+1}(\Psi^k)+V_j(\Psi^k)}(1-\xi) Z^{\flat}_j(\Psi^k, Y^k).$$

Therefore $Z^{\flat}_n(\Psi^k, Y^k) \leqslant e^{2\sum_j V_j(\Psi^k)}(1-\xi)^n Z^{\flat}_1(\Psi^k, Y^k)$. Since Lemma 7 implies $\sum_j V_j(\Psi^k) < \infty$, we have $P^{\flat}_G(\Psi^k) < \log(1-\xi) < 0$, as required. This completes the proof of the second claim. □

Now take $\theta_k > 0$ so that $P^{\flat}_G(\Psi^k + \theta_k \tau^k) \leqslant 0$. If the measure μ' from the beginning of the proof satisfies $h_{\mu'}(f) + \int \psi d\mu' > -\theta_k$, then $h_{\mu'}(f) + \int(\psi + \theta_k) d\mu' > 0$, so Lemma 5 implies that the corresponding induced measure μ'_{F_k} has $h_{\mu'_{F_k}}(F_k) + \int(\Psi^k + \theta_k \tau^k)\, d\mu'_{F_k} > 0$. From the Variational Principle for the system $(X^k_{\flat}, F_k, \Psi^k + \theta_k \tau^k)$ we see that μ'_{F_k} cannot be supported on type (ii) 1-cylinders of (X^k, F_k) only. Hence $\hat{\mu}'(\hat{X}^0) > 0$.

Finally take $\theta := \min\{\varepsilon_0, \theta_1, \dots, \theta_N\}$ and let μ be such that $h_\mu(f) + \int \psi d\mu > -\theta$. Since $\theta \leqslant \varepsilon_0$, we have $h_\mu(f) > \varepsilon$ by assumption, and therefore μ is compatible to (X^k, F_k) for some $k \in \{0, 1, \dots, N\}$. By the choice of θ and the argument of the previous paragraph, it follows that $\hat{\mu}(\hat{X}^0) > 0$ as required. □

5. Proofs of Theorem 1 and 2

Let $\varphi = \varphi_t = -t\log|Df|$, and Φ be the corresponding induced potential. Przytycki [Pr] proves that a measure $\mu \in \mathcal{M}$ is either supported on an attracting periodic orbit or $0 \leqslant \int \log|Df|\, d\mu < \infty$. So when we apply Lemma 5 to this potential, we will get finite integrals for both the measure on I and for the measure on the inducing scheme with the induced potential.

Lemma 7. *Assume that f has negative Schwarzian derivative. For inducing schemes obtained in Section 3, the induced potential has summable variations.*

Proof. In general, φ has unbounded variations. However, we note that inducing schemes as in Theorem 3 and Lemma 2 are maps $F : \bigcup_j X_j \to X$ with uniform Koebe space δ. Since φ is in general unbounded, it will not have bounded variations, but we only need to check that the induced potential Φ has bounded variations. By

the Koebe Lemma, $\frac{|DF(y)|}{|DF(x)|} < \frac{1+2\delta}{\delta^2} + 1$. Therefore

$$|\Phi(x) - \Phi(y)| = |t| \ \left| -\log|DF(x)| + \log|DF(y)| \right| = |t| \ \left| \log\left(\frac{|DF(y)|}{|DF(x)|}\right) \right|$$
$$\leqslant |t| \ \log\left(1 + \frac{1+2\delta}{\delta^2}\right) < |t| \left(\frac{1+2\delta}{\delta^2}\right).$$

By standard arguments, for any $\gamma > 1$ there exists $N = N(\gamma)$ such that we have $\inf_{x\in X}|DF^N(x)| > \gamma$ (here we use the negative Schwarzian assumption; alternatively a C^3 assumption and the absence of neutral periodic cycles would suffice). Moreover, F^N satisfies the above distortion estimates. Let $\gamma > \frac{1}{\delta}$ and let $G : \bigcup_j Y_j \to X$ be given by $G := F^N$ for $N = N(\gamma)$. Clearly, proving the lemma for Φ_N is sufficient.

We have that X is a $\gamma\delta$-scaled neighbourhood of Y_j for any j. Using the Koebe Lemma again for x, y in the same connected component of $G^{-1}(Y_j)$, we have

$$|\Phi_N(x) - \Phi_N(y)| < |t| \ \left(\frac{1+2\gamma\delta}{(\gamma\delta)^2}\right).$$

Repeating this argument for x, y in the same connected component of $G^{-n}(Y_j)$ that

$$|\Phi_N(x) - \Phi_N(y)| < |t| \ \left(\frac{1+2\gamma^n\delta}{(\gamma^n\delta)^2}\right) = |t|O(\gamma^{-n}).$$

Thus Φ_N, and hence Φ, has summable variations. □

The proofs of Theorems 1 and 2 have roughly the same structure. We start with the Collet-Eckmann case, leaving the additional details for the summable case to the end of the section. For use in both proofs, we define

$$Z_0(\Phi) := \sum_{F(x)=x} e^{\Phi(x)}.$$

As stated in the proof of Proposition 1, we have $Z_n(\Phi) = O(\|\mathcal{L}_\Phi 1\|_\infty^n)$. In this case, bounded distortion gives $\|\mathcal{L}_\Phi 1\|_\infty \asymp_{dis} Z_0(\Phi)$. Thus $Z_n(\Phi) = O([Z_0(\Phi)]^n)$.

We are now ready to prove Theorem 2, although we postpone the proof that $t \mapsto P(\varphi_t)$ is analytic to the end of Section 6.

Proof of the first part of Theorem 2. We choose X as in Corollary 1 and apply the method of Lemma 2 to get an extendible inducing scheme (X, F).

Fixing t, we define $\psi_S = \varphi_t - S$, and let Ψ_S be the induced potential. The natural candidate for S is $P(\varphi_t)$, but we will want to consider a more general value for this shift in the potential in order for (c) of Proposition 1 to hold.

We continue by showing that the induced system has bounded Gurevich pressure and (a) and (c) of Proposition 1 hold. As above, $Z_n(\Phi) = O(Z_0^n(\Phi))$. Therefore it suffices to show that $Z_0(\Phi_S) < \infty$ to conclude that $P_G(\Psi_S) < \infty$.

We wish to count the number of domains X_i with $\tau_i = n$. The number of *laps* of a piecewise continuous function g is the number of maximal intervals on which g is

monotone. We denote this number by laps(g). By [MSz], one characterisation of the topological entropy is $h_{top}(f) := \lim_{n\to\infty}\frac{1}{n}\log \mathrm{laps}(f^n)$. Therefore, for all $\varepsilon > 0$ there exists $C_\varepsilon > 0$ such that

$$\#\{\tau_i = n\} \leqslant \mathrm{laps}(f^n) \leqslant C_\varepsilon e^{n(h_{top}(f)+\varepsilon)}$$

for each n, where $h_{top}(f)$ denotes the topological entropy of f. Since f is Collet-Eckmann, the tail behaviour of the inducing scheme is exponential. This was shown for certain inducing schemes in [BLS]. We show in the proof of Proposition 3 that the results on the inducing schemes of [BLS] hold for the inducing schemes of Lemma 2. We also show there how [BRSS] allows us to strengthen the results of [BLS] to apply to maps with different critical orders, see Lemma 9 below.

For $t \leqslant 1$ we get

$$\begin{aligned}
Z_0(\Psi_S) &:= \sum_{F(x)=x} e^{\Psi_S(x)} = \sum_{i, x=F(x)\in X_i} e^{\Phi_t(x)-\tau_i(x)S} \\
&\asymp_{dis} \sum_i |X_i|^t e^{-\tau_i(x)S} = \sum_n \sum_{\tau_i=n} |X_i|^t e^{-nS} && \text{by the Koebe Lemma} \\
&\leqslant \sum_n \left(\sum_{\tau_i=n} |X_i|\right)^t e^{-nS} \left(\#\{\tau_i=n\}\right)^{1-t} && \text{by the Hölder inequality} \\
&\leqslant C_\varepsilon \sum_n e^{-\alpha n t} e^{-nS} e^{n(h_{top}(f)+\varepsilon)(1-t)} < \infty && \text{using tail behaviour}
\end{aligned}$$

provided t is sufficiently close to 1 and $S > h_{top}(f)(1-t) - \alpha t$. A similar estimate gives

$$\sum_i \tau_i e^{\Psi_S(x)} \asymp_{dis} \sum_i \tau_i |X_i|^t e^{-\tau_i S} < \infty. \tag{10}$$

For $t \geqslant 1$

$$\begin{aligned}
Z_0(\Psi_S) &\asymp_{dis} \sum_n \sum_{\tau_i=n} |X_i|^t e^{-nS} \leqslant \sum_n e^{-nS} \left(\sum_{\tau_i=n} |X_i|\right)^t \\
&\leqslant \sum_n e^{-\alpha n t} e^{-nS} < \infty,
\end{aligned}$$

provided $S > -\alpha t$. Similarly we can show

$$\sum_i \tau_i e^{\Psi_S(x)} \asymp_{dis} \sum_i \tau_i |X_i|^t e^{-\tau_i S} < \infty,$$

provided $S > -\alpha t$. When t is sufficiently close to 1, $P(\varphi_t)$ is close to 0, and thus if S is close to $P(\varphi_t)$ then the above sums are bounded.

Observe that the above estimates prove that condition (a) of Proposition 1 holds. For part (c) of that proposition, the estimates above prove that $P(\Psi_{P(\varphi_t)+\varepsilon}) < \infty$ for $\varepsilon < 0$ close to 0. Therefore, Corollary 1 shows that (c) is be satisfied. Therefore this inducing scheme gives rise to an equilibrium state $\mu_\varphi = \mu_\psi$. Moreover, from the proof of Proposition 1, $P_G(\Psi) = 0$.

It remains to show the uniqueness of the equilibrium state in $\mathcal{M}_+$, since up to this point we only know that μ_φ is the unique equilibrium state whose lift to the Hofbauer tower gives $\hat{X}$ positive mass. This follows from the next lemma.

Lemma 8. *If μ_φ is an equilibrium state, as above, compatible to an inducing scheme (X, F) then it is also is compatible to any other inducing scheme (X', F') provided $\hat{X}' \cap \mathcal{E} \neq \emptyset$. Here we assume that the inducing schemes are either both as in Theorem 3 or both as in Lemma 2.*

Proof. We will assume that the inducing schemes here are all as in Lemma 2, since this is the more difficult case. Let $(\hat{X}, \hat{F})$ be the inducing scheme used above. The proof follows if we can show that $\hat{\mu}_\varphi(\hat{X}') > 0$.

Transitivity of $(\mathcal{E}, \hat{f})$ implies that there exists $n \geqslant 0$ so that $\hat{f}^{-n}(\hat{X}') \cap \hat{X}$ contains an open set. As in Proposition 2, since μ_Ψ gives positive mass to cylinders, this implies that there exists $\hat{U} \subset \hat{X}$ so that $\hat{\mu}_\varphi(\hat{U}) > 0$ and $\hat{f}^n(\hat{U}) \subset \hat{X}'$. Hence,

$$\hat{\mu}_\varphi(\hat{X}') \geqslant \hat{\mu}_\varphi(\hat{f}^n(\hat{U})) \geqslant \hat{\mu}_\varphi(\hat{U}) > 0.$$

Therefore, μ_φ is compatible to (X', F'). □

Suppose that $\mu \in \mathcal{M}_+$ is an equilibrium state. By the ideas of Lemma 2 there must exist a first extendible inducing scheme (X', F', Ψ') which is compatible to μ and which corresponds to a first return map to a set $\hat{X}'$ on the Hofbauer tower. Lemma 8 implies that μ_φ is compatible to (X', F') and hence $\mu = \mu_\varphi$ by the uniqueness of equilibrium states on an inducing scheme. □

To do the summable case, we adapt techniques from [BLS]. In that paper, the Bounded Backward Contraction is used for arbitrary neighbourhoods of the critical set, which at the time was only known to hold when all critical orders ℓ_c are the same. Using results from [BRSS], and specifying the neighbourhoods U, we can improve this in the following lemma.

Lemma 9. *Let $f \in \mathcal{H}$ be a multimodal map with negative Schwarzian derivative such that $\lim_{n\to\infty} |Df^n(f(c))| = \infty$ for each $c \in \mathrm{Crit}$. Then for any $\varepsilon > 0$ and $\lambda > 1$, we can find critical neighbourhoods $U := f^{-1}(B_\varepsilon(f(\mathrm{Crit})))$ that are* λ-nice *in the sense that*

- *$f^n(\partial U) \cap U = \emptyset$ for all $n \geqslant 0$;*
- *if $V \subset U$ is the domain of the first return map to U, then the interval V' concentric to V and of length $(1 + 2\lambda)|V|$ is contained in U.*

Moreover, there exists $b > 0$ such that

$$|Df^r(x)| \geqslant b \text{ for all } x \in I \text{ and } r = \min\{n \geqslant 0 : f^n(x) \in U\}, \tag{11}$$

where the λ-nice critical neighbourhood U can be chosen arbitrarily small.

Proof. The first part follows immediately from [BRSS] which considers C^3 non-flat multimodal maps. Our assumption that f is C^2 with negative Schwarzian derivative

actually gives a slightly stronger version of the Koebe distortion theorem, and hence is sufficient to claim the results from [BRSS]. Lemma 3 in [BRSS] shows the existence of λ-nice neighbourhoods U of Crit. Denote the connected components of U by U^c, $c \in$ Crit. If $r = r(x) \geqslant 0$ is the first entrance time of x to U, then the niceness of U guarantees that there exists an interval J_x so that f^r maps J diffeomorphically onto U^c for some $c \in$ Crit. If $f^r(x)$ belongs to first return domain V, then there is $J_V \subset J$ such that $f^r : J_V \to V$ is monotone with distortion bound depending only on λ. A special case of this is when $V := \tilde{U}^c$ is the central return domain in U^c. Let $\tilde{U} = \cup_{c\in\mathrm{Crit}}\tilde{U}^c$. In this case, the first entrance time $\tilde{r} \geqslant 0$ of any x into $\tilde{U}$ corresponds to a diffeomorphic branch $f^{\tilde{r}} : \tilde{J} \to \tilde{U}^c$ with distortion bound depending only on λ.

Remark 7. *Note that $U \subset f^{-1}(B_\varepsilon(f(\mathrm{Crit}))$, where ε can be taken arbitrarily small. As a result, the components U^c need not have comparable sizes for all $c \in$ Crit, but scale as ε^{1/ℓ_c}. A similar difference in size is true for the components of $\tilde{U}$, and this is a major difference with the critical neighbourhoods as used in* [BLS]. *If all components of $\tilde{U}$ have the same size, then* (11) *can fail.*

To prove (11), fix a λ-nice critical neighbourhood U_0, and let $U_1 := \tilde{U}_0$ be the union of its central return domains. This set is λ-nice again. There exists $b = b(U_1) > 0$ such that for every $x \in I$, $|Df^{r_1}(x)| \geqslant b$ for $r_1 = \min\{n \geqslant 0 : f^n(x) \in U_1\}$. Continue to construct λ-nice neighbourhoods $U_i = \tilde{U}_{i-1}$ as the union of the central return domains of the previous stage. These set shrink at least exponentially in i, so we obtain a λ-nice neighbourhood $U = U_p$ as small as we want.

Now let $r_1 \leqslant r_2 \leqslant \ldots \leqslant r_p = r$ be the return times of x to $U_1 \supset U_2 \supset \cdots \supset U_p$. There is a neighbourhood $J \ni x$ such f^r maps J diffeomorphically onto a component of U. The maps $f^{r_{i+1}-r_i}|_{f^{r_i}(J)}$ are composition of monotone branches of the first return map to U_i. If λ is sufficiently large, then these branches are expanding, uniformly in x. Hence $|Df^r(x)| \geqslant |Df^{r_1}(x)| \geqslant b$. □

Proposition 3. *Suppose that f is a multimodal map satisfying* (1)*. Then on every sufficiently small cylinder set X there is a first extendible return inducing scheme (X, F, τ) and $t_1 \in [t_0, 1]$ such that for all $t \in (t_1, 1]$: and all potential shifts $S \geqslant 0$:*

$$Z_0(\Psi_S) := \sum_{F(x)=x} e^{\Psi_S(x)} < \infty,$$

where Ψ_S is the induced potential of the shifted potential $\psi_S := \varphi_t - S$. Furthermore for the equilibrium state $\mu_{\Psi_{P(\varphi_t)}}$, $\mu_{\Psi_{P(\varphi_t)}}\{\tau = n\}$ decays exponentially for $t \in (t_1, 1)$, and polynomially for $t = 1$.

Proof. For the case $t = 1$, if the critical points all have the same order then [BLS] gives an inducing scheme with polynomial tails (this is also sufficient to show $Z_0(\Psi_S) < \infty$ for all $S \geqslant 0$). Below we show that inducing schemes from Lemma 2 fit into the framework of [BLS]. We also show that by Lemma 9, the machinery of [BLS] can also be applied to maps with critical points with different critical orders, by Lemma 9. We focus on the details of the case $t < 1$, showing that these systems have exponential tails. The proof that our inducing schemes give equilibrium states

with polynomial tails for $t = 1$ is left to the reader. From here onwards, we restrict our proof to the case $t < 1$.

Fix a single cylinder set $X \in \mathcal{P}_n$ and $\delta \in (0, \frac{1}{2})$ so small that a δ-scaled neighbourhood of X is contained in $\pi(D)$ for at least one domain D of the closed primitive subgraph $\mathcal{E}$ (cf. Lemma 1) of the Hofbauer tower. The inducing scheme will be the first extendible return to X in the sense of Lemma 2: namely, for each X_i, there is a neighbourhood X_i' such that f^{τ_i} maps X_i' diffeomorphically onto a δ-scaled neighbourhood X. Let $\hat{X} \subset \pi^{-1}(X)$ be such that the inducing scheme corresponds to the first return map to $\hat{X}$. Since X is a cylinder set, $\hat{X}$ is *nice* in the sense that for $n \geqslant 1$, $\hat{f}^n(\hat{x})$ never intersects the interior of $\hat{X}$ for each $\hat{x} \in \partial\hat{X}$. There is a dense orbit $\mathrm{orb}(\hat{y})$ in $\mathcal{E}$, and for each visit $\hat{y}' \in \mathrm{orb}(\hat{y}) \cap \hat{X}$, there is a neighbourhood $\hat{X}_i \ni \hat{y}'$ such that $\hat{f}^{\tau_i} : \hat{X}_i \to \hat{X}$ is extendible to a δ-scaled neighbourhood of a component of $\hat{X}$. Therefore, the union $\cup_i X_i$ (and hence X^∞) is dense in X, and the niceness of $\hat{X}$ guarantees that the sets X_i are pairwise disjoint.

Note that (1) implies that

$$\sum_n \left(\gamma_n^{\ell_c - 1} |Df^n(c_1)|\right)^{-t_0/\ell_c} < \infty, \tag{12}$$

for every $c \in \mathrm{Crit}$, some $t_0 < 1$ and summable sequence $\{\gamma_n\}_{n\in\mathbb{N}}$ with $\gamma_n \in (0, \delta|X|)$. Throughout we can take $\gamma_n = \frac{\delta|X|}{n \log^2(n+10)}$.

We use ideas and results of [BLS] extensively. To start with, given a neighbourhood U of Crit as in Lemma 9 (so that (11) holds), we can assign to any $x \in I$ a sequence of *binding periods* along which the orbit of x shadows a critical orbit, followed by *free period* during which the orbit of x remains outside U. During the binding period, derivative growth is comparable to derivative growth of the critical orbit. The precise definition of binding period of $x \in U$ is:

$$p(x) = \min\{k \geqslant 1 : |f^k(x) - f^k(c)| \geqslant \gamma_k |f^k(c) - \mathrm{Crit}|\},$$

where c is the critical point closest to x. At the end of the binding period, derivatives have recovered from the small derivative incurred close to c. Indeed, Lemma 2.5 of [BLS] states that there is $C_0 > 0$, independent of U, such that

$$F_p'(x) := \inf\{|Df^p(x)| : x \in U, p(x) = p\} \geqslant C_0 \left(\gamma_p^{\ell_c - 1} |Df^p(f(c))|\right)^{1/\ell_c}.$$

where c is the critical point closest to x. If U is a small neighbourhood, then $p(x)$ is big. Hence we can take U so small that the minimal binding period $p_U := \min\{p(x) : x \in U\}$ is so large that Equation (5) in [BLS] holds:[1]

$$\max_{c \in \mathrm{Crit}} \sum_{s \leqslant n} \sum_{\substack{(p_1, \dots, p_s) \\ \sum_i p_i \leqslant n \\ p_i \geqslant p_U}} \prod_{p_i} \zeta \left(\gamma_{p_i}^{\ell_c - 1} |Df^{p_i}(f(c))| \right)^{-1/\ell_c} \leqslant 1. \tag{13}$$

[1]Here we take into account the typo in Equation (5) of [BLS] where the $-$ in the exponent is missing.

Here $\zeta = 4C_4\#\mathrm{Crit}$ (see later in the proof) is a fixed number involving a Koebe constant and a constant emerging from the Bounded Backward Contraction Condition (11), see Lemma 9. The constant ζ is independent of U.

During the *free period*, derivatives grow exponentially (Mañé's Theorem, see [MSt, Theorem III.5.1.]), because there exist $C_1 > 0$ and $\lambda_1 > 1$, depending only on f and U, such that

$$|Df^k(x)| \geqslant C_1\lambda_1^k \quad \text{if} \quad f^i(x) \notin U \text{ for } 0 \leqslant i < k. \tag{14}$$

Now fix a neighbourhood U of Crit as in Lemma 9 with $\partial U \subset \cup_n f^{-n}(\mathrm{Crit})$ and so small that estimate (13) holds. In fact, parallel to (14), one can derive sets that avoid U for a long time are exponentially small: there are $C_a > 0$ and $\lambda_2 > 1$ such that

$$|f^n(A)| \leqslant C_a\lambda_2^{-k} \quad \text{if} \quad f^i(A) \cap U = \emptyset \text{ for } 0 \leqslant i < k. \tag{15}$$

Since ∂U consists of precritical points, and each X_i is mapped monotonically onto X, there is κ such that $f^j(X_i) \cap \partial U \neq \emptyset$ implies $j \geqslant \tau_i - \kappa$. Given X_i and $j < \tau_i - \kappa$, $f^j(X_i)$ will either be contained in or disjoint from U. Thus we can define $\nu_j(X_i)$ to be the time at which the j-th binding period starts and the binding periods itself as $p_j(X_i) = \min\{p_j(x) : x \in X_i\}$. Since $f^{\tau_i - n}$ maps $f^n(X_i)$ to X in an extendible way for each $n \leqslant \tau_i$, the distortion of $f^{\tau_i - n}|_{f^n(X_i)}$ is bounded uniformly in i and n. We will write $\nu_j = \nu_j(X_i)$ and $p_k = p_k(X_i)$ if it is clear from the context which X_i is meant. Note that the inducing time τ_i of X_i cannot be inside a binding period, because during the binding period, X_i shadows some critical value $f^k(c)$ γ_k-closely, and $\gamma_k < \delta|X|$ for every k.

In the terminology of [BLS], every return time is a *deep return*, and there are no *shallow returns*. Let τ_i' be the time that the final binding period ends, so $\tau_i' = \nu_s + p_s \leqslant \tau_i$ if X_i has s binding periods.

To estimate $Z_0(\Psi_S)$, we first group together domains X_i into a 'cluster' if they have the same binding periods $p_1, \dots, p_s$ up to their common time τ_i' and $f^j(\mathrm{conv}\ \tilde{A}) \cap \mathrm{Crit} = \emptyset$ for $j \leqslant \tau_i$, where conv $\tilde{A}$ is the convex hull of the cluster. We have by the Hölder inequality

$$\begin{aligned}
Z_0(\Psi_S) &\asymp_{dis} \sum_i |X_i|^t e^{-\tau_i S} = \sum_n e^{-nS} \sum_{n' \leqslant n} \sum_{\substack{\text{cluster } \tilde{A} \\ \tau(\tilde{A})=n, \tau'(\tilde{A})=n'}} \sum_{X_i \subset \tilde{A}} |X_i|^t \\
&\leqslant \sum_n e^{-nS} \sum_{n' \leqslant n} \sum_{\substack{\text{cluster } \tilde{A} \\ \tau(\tilde{A})=n, \tau'(\tilde{A})=n'}} (\#\{i : X_i \text{ belongs to } \tilde{A}\})^{1-t} \left(\sum_{X_i \subset \tilde{A}} |X_i| \right)^t \\
&\leqslant \sum_n e^{-nS} \sum_{n' \leqslant n} e^{(h_{top}(f)+\varepsilon)(n-n')(1-t)} \sum_{\substack{\text{cluster } \tilde{A} \\ \tau(\tilde{A})=n,\ \tau'(\tilde{A})=n'}} |\mathrm{conv}\tilde{A}|^t,
\end{aligned}$$

where the cardinality $\#\{i : X_i \text{ belongs to } \tilde{A}\}$ is estimated by $e^{(h_{top}(f)+\varepsilon)(n-n')}$ for some small $\varepsilon = \varepsilon(t) > 0$, because the cluster $\tilde{A}$ has $n - n'$ iterates left to the inducing time.

To estimate $\sum_{\tau(\tilde{A})=n,\tau'(\tilde{A})=n'}|\tilde{A}|^t$, we distinguish two classes of clusters depending on the amount of free time in the first τ' iterates. For $\eta > 0$ to be fixed later, and for given n and n', let

$$\hat{\mathcal{P}}'_{n,n'} = \left\{ \tilde{A} \ : \ \tau'(\tilde{A}) = n', \tau(\tilde{A}) = n, \sum_{i=1}^{s} p_i \leqslant \eta n \right\}$$

and

$$\hat{\mathcal{P}}''_{n,n'} = \left\{ \tilde{A} \ : \ \tau'(\tilde{A}) = n', \tau(\tilde{A}) = n, \sum_{i=1}^{s} p_i > \eta n \right\}.$$

The estimates for $\hat{\mathcal{P}}'_{n,n'}$ and $\hat{\mathcal{P}}''_{n,n'}$ will use Lemmas 3.5 and 3.6 of [BLS] respectively. Indeed, Lemma 3.5 of [BLS] gives some η (fixing the definition of $\hat{\mathcal{P}}'_{n,n'}$) and $\lambda_3 > 1$ depending on λ_1 and η such that

$$\sum_{\tilde{A}\in\hat{\mathcal{P}}'_{n,n'}} |\tilde{A}|^t \leqslant \lambda_3^{-\frac{1}{2}n't} \sup_{\tilde{A}\in\hat{\mathcal{P}}'_{n,n'}} |f^{n'}(\tilde{A})|^t \leqslant C_1^{-t}\lambda_3^{-\frac{1}{2}n't}\lambda_1^{-(n-n')t}, \tag{16}$$

where the last inequality follows by (14) because $f^{n'}(\tilde{A})$ is disjoint from U for the remaining $n - n'$ iterates.

Continuing with this η, define $d_n(c) := \min_{i<n}(\gamma_i/|Df^i(f(c))|)^{1/\ell_c}|f^i(c) - \mathrm{Crit}| \leqslant 1$ (formula (2) in [BLS]) and let (following [BLS, page 635])

$$\hat{d}_{n,j}(c) = d_i(c) \text{ for } i = \max\left\{ \left\lceil \frac{\eta n}{2j^2} \right\rceil, 1 \right\}.$$

Then an adaptation of Lemma 3.6 of [BLS] gives a constant $C_2 > 0$ such that

$$\sum_{\tilde{A}\in\hat{\mathcal{P}}''_{n,n'}} |\tilde{A}|^t \leqslant C_1^{-t}\lambda_1^{-(n-n')t} C_2 \sum_{s=1}^{n'} 2^{-j} \left(\max_{c\in\mathrm{Crit}} \hat{d}_{n',j} \right)^t. \tag{17}$$

Indeed, select the longest binding period among $(p_1, \dots, p_s)$ of the cluster, and call it p_j. Note that $p_j > \eta n/(2j^2)$, because otherwise $\sum_{k=1}^{s} p_k < \eta n$, contradicting the definition of $\hat{\mathcal{P}}''_{n,n'}$. The interval $[x, y] := f^{\nu_j}(\mathrm{conv}\ \tilde{A})$ satisfies

$$|x - y| \leqslant C_3 \max_{p\geqslant \eta n/2j^2} d_p(c) \cdot |f^{\nu_j+p_j}(\mathrm{conv}\ \tilde{A})| = C_3\ \hat{d}_{n',j}(c) \cdot |f^{\nu_j+p_j}(\mathrm{conv}\ \tilde{A})|,$$

where C_3 is a uniform distortion constant. Write $\tilde{A} = \tilde{A}_{p_1,\dots,p_j}$ to indicate that p_j is the longest binding period of $\tilde{A}$. By Lemma 3.2 of [BLS], and recalling that all returns are *deep*, we can find C_4 such that

$$|\tilde{A}_{p_1,\dots,p_j}| \leqslant C_4^{j-1}\ |f^{\nu_{j-1}+p_{j-1}}(\mathrm{conv}\ \tilde{A}_{p_1,\dots,p_j})| \prod_{k=1}^{j-1} \frac{1}{F'_{p_k}}.$$

Following the proof of Lemma 3.6 of [BLS], we obtain

$$\begin{aligned}\sum_{\substack{\text{cluster } \tilde{A} \\ \tau(\tilde{A})=n,\tau'(\tilde{A})=n'}} |\tilde{A}|^t &\leqslant \sum_{j=1}^{n'} \sum_{(p_1,\dots,p_j)} |\tilde{A}_{p_1,\dots,p_j}|^t \\ &\leqslant \sum_{j=1}^{n'} \left(C_3 \max_{c\in \text{Crit}} \hat{d}_{n',j}(c)\right)^t \\ &\qquad \times \sum_{(p_1,\dots,p_j)} (2\#\text{Crit})^j \left(C_4^{j-1} \prod_{k=1}^{j-1} \frac{1}{F'_{p_k}} \right)^t |f^{\nu_j+p_j}(\text{conv } \tilde{A}_{p_1,\dots,p_j})|^t,\end{aligned}$$

where the $(2\#\text{Crit})^j$ accounts for the different sides of critical points that have intervals with the same binding period. Using (13) with $\zeta = 4C_4\#\text{Crit}$, we can estimate this by

$$\sum_{j=1}^{n'} \left(C_3 \max_{c\in \text{Crit}} \hat{d}_{n',j}(c)\right)^t \cdot 2^{-j} \cdot |f^{\nu_j+p_j}(\widetilde{\text{conv}}\ A_{p_1,\dots,p_j})|^t.$$

The maps $f^{\nu_j+p_j}|_{\text{conv } \tilde{A}_{p_1,\dots,p_j}}$ and $f^{n'-(\nu_j+p_j)}|_{f^{\nu_j+p_j}(\text{conv } \tilde{A}_{p_1,\dots,p_j})}$ have bounded distortion. Each set $f^{n'}(\tilde{A}_{p_1,\dots,p_j})$ is disjoint from U for the remaining $n-n'$ iterates, so (15) gives $|f^{n'}(X_i)| \leqslant C_1^{-1}\lambda_2^{-(n-n')}$. Therefore

$$\sum_{\substack{\text{cluster } \tilde{A} \\ \tau(\tilde{A})=n,\tau'(\tilde{A})=n'}} |\tilde{A}|^t \ \leqslant C_1^{-t}\lambda_2^{-(n-n')t} C_2 \sum_{j=1}^{n'} 2^{-j} \left(\max_{c\in \text{Crit}} \hat{d}_{n',j}(c)\right)^t,$$

for $C_2 = (C_3C_4)^t$. This proves (17).

Now we obtain (using (17) and (16))

$$\begin{aligned}Z_0(\Psi_S) &\leqslant \sum_n e^{-nS} \sum_{n'\leqslant n} e^{(h_{top}(f)+\varepsilon)(n-n')(1-t)} \left(\sum_{\tilde{A}\in \hat{\mathcal{P}}'_{n,n'}} |\tilde{A}|^t + \sum_{\tilde{A}\in \hat{\mathcal{P}}''_{n,n'}} |\tilde{A}|^t \right) \\ &\leqslant \sum_n e^{-nS} \sum_{n'\leqslant n} e^{(h_{top}(f)+\varepsilon)(n-n')(1-t)} \lambda_2^{-(n-n')t} \left(\lambda_3^{-\frac{1}{2}n't} + \sum_{j=1}^{n'} 2^{-j} \left(\max_{c\in \text{Crit}} \hat{d}_{n',j}(c)\right)^t \right),\end{aligned}$$

which is finite, provided t is sufficiently close to 1. The proof that $\int \tau \ d\mu_\Psi < \infty$ amounts to showing that $ne^{-nS}\sum_{n'\leqslant n}\sum_{\tau_i=n,\tau'_i=n'} |X_i|^t$ is summable in n, cf. (10). If $t<1$, then $S = P(\varphi) > 0$ by (2), so for t sufficiently close to 1, the exponential factor e^{-nS} dominates n and summability follows. This also implies the required exponential tails property for $(X, F, \mu_{\Psi_{P(\varphi_t)}})$. □

For the case $t=1$ we already know by [BRSS] that there is an acip, so the above proposition shows that the acip must have polynomial tails. Hence the proof of Theorem 1 for (except for the proof that $t\mapsto P(\varphi_t)$ is analytic, which is postponed to the end of Section 6) essentially amounts to an application of Proposition 1

(Case 4.) to the case $t \in (t_1, 1)$, and is completed in a similar way to the proof of Theorem 2. The rate of decay of the tails follows from Proposition 3.

The following lemma, which will be particularly useful in Section 6, implies that we can fix an inducing scheme so that any measure with large free energy, for some φ_t, must be compatible to this inducing scheme.

Lemma 10. *For any point $x \in I$ there exists an inducing scheme (X, F) as in Lemma 2 with $x \in X$ and so that the following hold.*

- *In the case of, and with $t_1 < 1$ as in Theorem 1 (polynomial growth rate): for any $t_1 < t_2 < 1$ there exists $\varepsilon_0 > 0$ so that for all $t \in (t_1, t_2)$, if $h_\mu(f) + \int \psi_t \, d\mu > P_+(\psi_t) - \varepsilon_0$ then μ is compatible to (X, F).*
- *In the case of, and with $t_1 < 1 < t_2$ as in Theorem 2 (Collet-Eckmann): there exists $\varepsilon_0 > 0$ so that for all $t \in (t_1, t_2)$, if $h_\mu(f) + \int \psi_t \, d\mu > P_+(\psi_t) - \varepsilon_0$ then μ is compatible to (X, F).*

Proof. By Lemma 4, there exist $\varepsilon_0, \varepsilon > 0$ such that for any $t \in (t_1, t_2)$, $h_\mu(f) + \int \psi_t \, d\mu > P_+(\psi_t) - \varepsilon_0$ implies $h_\mu(f) > \varepsilon$. We can choose $\{\hat{X}^k\}_{0 \leqslant k \leqslant N}$ as in Proposition 2: we need only select these sets so small that the corresponding inducing scheme is uniformly expanding, in order to satisfy (a) of that lemma, and so that $x \in \pi(\hat{X}^0)$. Property (b) of Proposition 2 follows for all $t \in (t_1, t_2)$ by the computations in the proof of Theorem 2 and in Proposition 3. The fact that for any $t \in (t_1, t_2)$, μ_t is compatible to our (X^0, F_0) follows by Lemma 8. Therefore, Proposition 2 implies that the measures μ must be compatible to (X^0, F_0). Finally take $(X, F) = (X^0, F_0)$. □

6. Exponential Tails and Positive Discriminant

In Theorems 1 and 2 we see that with the exception of non-Collet-Eckmann maps (i.e., satisfying (1) but not (3)) with potential $\varphi = -\log |Df|$, all the equilibrium states μ_φ obtained are compatible to an inducing scheme with exponential tail behaviour: $\mu_\Psi(\{x \in X : \tau(x) = n\}) \leqslant Ce^{-\alpha n}$ for some $C, \alpha > 0$.

The literature gives many consequences; we mention a few:

- The system (I, f, μ_φ) has exponential decay of correlations and satisfies the Central Limit Theorem. This follows directly from Young's results [Y] relating the decay of correlations to the tail behaviour of the Young tower.
- The system (I, f, μ_φ) satisfies the Almost Sure Invariance Principle (ASIP), see [MN] or [HK1] for earlier ideas in this direction.
- In [C], Collet proves Gumbel's Law (which is related to exponential return statistics) for the acip provided the Young tower construction has exponential tail behaviour. It seem likely that this result extends to the equilibrium states for $\varphi_t = -t \log |Df|$ and $t < 1$.

Another application of exponential tails pertains to analyticity of the pressure function $t \mapsto P(\varphi_t)$ and the absence of phase transitions (which would be expressed by

lack of differentiability of the pressure function). A key result here is phrased by Sarig [Sa2] in terms of directional derivatives

$$\frac{d}{ds}P(\psi + s\upsilon)|_{s=0}$$

where ψ and υ are suitable potentials. To prove analyticity of $t \mapsto P(t\varphi)$ near $t = 1$, we take $\upsilon = \psi = \varphi$. Sarig obtains his results for Gurevich pressure. For appropriate potentials and inducing scheme, he first introduces the concept of discriminant $\mathfrak{D}$, which is positive if and only if the inducing scheme has exponential tails with respect to the equilibrium state of the induced potential. Next it is shown that if the inducing scheme is a first return map, then positive discriminant implies analyticity of $s \mapsto P_G(\psi + s\upsilon)$ near $s = 0$. In our case, the inducing scheme is a first return map on the Hofbauer tower, but also a Rokhlin-Kakutani tower can be constructed for which the first return map to the base is isomorphic to the inducing scheme. Currently, in the context of smooth dynamical systems, these towers tend to be called a Young towers [Y]. It is the better distortion properties than the Young tower on elements of its natural partition $\Delta_{i,j}$, see below, that makes us prefer the Young tower over the Hofbauer tower in the section.

The resulting analyticity of the pressure function on the Young tower then needs to be related to the original system. We will do that using a transition from Gurevich pressure to the following type of pressure:

$$P_+(\psi) := \sup\left\{h_\mu(f) + \int \psi \, d\mu : \mu \in \mathcal{M}_+ \text{ and } -\int \psi \, d\mu < \infty\right\}$$

for which we use a result by Fiebig et al. [FFY].

The set-up of the remainder of this section is as follows. We first introduce the Young tower associated with the inducing scheme, and then discuss directional derivatives and discriminants. This gives us the necessary terminology to state the main theorem (Theorem 5). Then we show how this can be applied to prove the remaining analyticity parts of Theorems 1 and 2. Finally, we prove Theorem 5.

Let $X \subset I$ and (X, F, τ) be an inducing scheme on X where $F = f^\tau$. As usual we denote the set of domains of the inducing scheme by $\{X_i\}_{i\in\mathbb{N}}$. The *Young tower*, see [Y], is defined as the disjoint union

$$\Delta = \bigsqcup_{i\in\mathbb{N}} \bigsqcup_{j=0}^{\tau_i - 1} (X_i, j),$$

with dynamics

$$f_\Delta(x,j) = \begin{cases} (x, j+1) & \text{if } x \in X_i, j < \tau_i - 1; \\ (F(x), 0) & \text{if } x \in X_i, j = \tau_i - 1. \end{cases}$$

For $i \in \mathbb{N}$ and $0 \leqslant j < \tau_i$, let $\Delta_{i,j} := \{(x,j) : x \in X_i\}$ and $\Delta_l := \bigcup_{i\in\mathbb{N}} \Delta_{i,l}$ is called the *l-th floor*. Define the natural projection $\pi_\Delta : \Delta \to X$ by $\pi_\Delta(x,j) = f^j(x)$, and $\pi_X : \Delta \to X$ by $\pi_X(x,j) = x$. Note that (Δ, f_Δ) is a Markov system, and the first return map of f_Δ to the *base* Δ_0 is isomorphic (X, F, τ).

Also, given $\psi : I \to \mathbb{R}$, let $\psi_\Delta : \Delta \to \mathbb{R}$ be defined by $\psi_\Delta(x,j) = \psi(f^j(x))$. Then the induced potential of ψ_Δ to the first return map to Δ_0 is exactly the same as the induced potential of ψ to the inducing scheme (X, F, τ).

The differentiability of the pressure functional can be expressed using directional derivatives $\frac{d}{ds} P_G(\psi + s\upsilon)\Big|_{s=0}$. We will use the method of [Sa2], but will require less stringent conditions on the potentials. Let (W, f) be a topologically mixing dynamical system with the set of n-cylinders denoted by $\mathcal{Q}_n$. For a potential $\psi : W \to [-\infty, \infty]$ we can ask that ψ satisfies

$$\sup_{C_n \in \mathcal{P}_n} \sup_{x,y \in C_n} |\psi_n(x) - \psi_n(y)| = o(n). \tag{18}$$

As shown in [FFY], this guarantees that ψ satisfies (9) which means that the Gurevich pressure is well defined and independent of the initial cylinder set X_i, where $Z_n(\psi) = Z_n(\psi, X_i)$; also Theorem 7 below is satisfied. Moreover, if the induced potential is weakly Hölder continuous, then (18) is a sufficient condition on the original potential to allow us to use the results of [Sa2, Section 6], see Theorem 6.

For an inducing scheme (X, F, τ), let ψ_Δ and υ_Δ be the lifted potentials to the Young tower. Suppose that $\psi_\Delta : \Delta \to \mathbb{R}$ satisfies (18). We define the set of *directions* with respect to ψ as the set

$$Dir_F(\psi) := \left\{ \upsilon : \sup_{\mu \in \mathcal{M}_+} \left| \int \upsilon \, d\mu \right| < \infty, \upsilon_\Delta \text{ satisfies (18)}, \sum_{n=2}^{\infty} V_n(\Upsilon) < \infty, \text{ and} \right.$$
$$\left. \exists \varepsilon > 0 \text{ s.t. } P_G(\psi_\Delta + s\upsilon_\Delta) < \infty \ \forall \ s \in (-\varepsilon, \varepsilon) \right\},$$

where Υ is the induced potential of υ. As in previous sections, let $\psi_S := \psi - S$ (and so $\Psi_S = \Psi - S\tau$). Set $p_F^*[\psi] := \inf\{S : P_G(\Psi_S) < \infty\}$.[2] If $p_F^*[\psi] > -\infty$, we define the X-*discriminant* of ψ as

$$\mathfrak{D}_F[\psi] := \sup\{P_G(\Psi_S) : S > p_F^*[\psi]\} \leqslant \infty.$$

Given a dynamical system (X, F), we say that a potential $\Psi : X \to \mathbb{R}$ is *weakly Hölder continuous* if there exist $C, \gamma > 0$ such that $V_n(\Psi) \leqslant C\gamma^n$ for all $n \geqslant 0$.

The main result of this section is as follows:

Theorem 5. *Let $f \in \mathcal{H}$ be an interval map with potential $\varphi : I \to (-\infty, \infty]$. Suppose that φ satisfies* (18) *or is of the form $\varphi = -t \log |Df|$. Take $\psi = \varphi - P(\varphi)$. Then $\mathfrak{D}_F[\psi] > 0$ if and only if (X, F, μ_Ψ) has exponential tails.*

Moreover, the inducing scheme can be chosen such that given $\upsilon \in Dir_F(\psi)$ such that $\psi_\Delta + \upsilon_\Delta$ is continuous and the induced potential Υ is weakly Hölder continuous, there exists $\varepsilon > 0$ such that $s \mapsto P_+(\psi + s\upsilon)$ is real analytic on $(-\varepsilon, \varepsilon)$.

As noted before, the appropriately shifted potential $\varphi_t = -t \log |Df|$, gives rise to an equilibrium state with exponential tail for t in a neighbourhood of 1 if (3) holds, and for $t \in (t_1, 1)$ if (3) fails but (1) holds. Take $\upsilon = -\log |Df|$. Any induced system

[2] Note that we use the opposite sign for $p_F^*[\psi]$ to Sarig.

provided in Section 5 is extendible, so by the Koebe lemma the induced potential Υ has summable variations, and in fact is weakly Hölder. Similarly $(-\log|Df|)_\Delta$ satisfies (18). Also, since $P_G(\psi_\Delta + s\upsilon_\Delta) \leqslant P_G(\Psi + s\Upsilon)$ which is clearly bounded for small s, we have the $P_G(\psi_\Delta + s\upsilon_\Delta) < \infty$ for small s. Therefore there is an inducing scheme with $\upsilon \in Dir_F(\psi)$. Thus Theorem 5 can be applied to give the analyticity of $t \mapsto P(\varphi_t)$ for $t \in (t_1, 0)$, to complete the proofs of Theorems 1 and 2.

Proof. Suppose that $\mathfrak{D}_F[\psi] > 0$. This is equivalent to the existence of $0 > \varepsilon_0 > p_F^*[\psi]$ such that $P_G(\Psi_{\varepsilon_0}) < \infty$. By the Gibbs property, for $\varepsilon > \varepsilon_0$ we have $\mu_{\Psi_\varepsilon}(\{\tau = n\}) \asymp \sum_{\tau_i = n} e^{\Psi_i - n\varepsilon}$. Then

$$\mu_{\Psi_\varepsilon}(\{\tau = n\}) \asymp e^{-n(\varepsilon-\varepsilon_0)} \sum_{\tau_i = n} e^{\Psi_i - n\varepsilon_0}.$$

Notice that

$$\sum_{\tau_i = n} e^{\Psi_i - n\varepsilon_0} \asymp \mu_{\Psi_{\varepsilon_0}}(\{\tau = n\}) < \mu_{\Psi_{\varepsilon_0}}(X) = 1,$$

so $\mu_{\Psi_\varepsilon}(\{\tau = n\}) < Ce^{-n(\varepsilon-\varepsilon_0)}$. Since $\varepsilon - \varepsilon_0 > 0$, $(X, F, \mu_{\Psi_\varepsilon})$ has exponential tails.

Conversely, suppose that (X, F, μ_Ψ) has exponential tails with exponent $\alpha > 0$, that is

$$\sum_{\tau_i = n} e^{\Psi_i} \asymp \mu_\Psi(\{\tau = n\}) < Ce^{-n\alpha}.$$

Then, for all $-\alpha < \varepsilon_0$, and for Z_0 defined on page 20,

$$P_G(\Psi_{\varepsilon_0}) \leqslant CZ_0(\Psi_{\varepsilon_0}) \leqslant C\sum_n \sum_{\tau_i = n} e^{\Psi_i - n\varepsilon_0} < C\sum_n e^{-n(\alpha+\varepsilon_0)} < \infty.$$

Therefore $p_F^*[\psi] \leqslant -\alpha < 0$ and so $\mathfrak{D}_F[\psi] > 0$.

For the second part of the theorem, we use the following result from [Sa2, Theorem 4].

Theorem 6. *Let (W, f) be a topologically mixing dynamical system and $\psi : W \to (-\infty, \infty]$ be a potential satisfying* (18)*, such that $P_G(\psi) < \infty$ and for $X \in \mathcal{P}_n$, $\mathfrak{D}_F[\psi] > 0$ and Ψ is weakly Hölder continuous. Then for all $\upsilon \in Dir_F(\psi)$ such that Υ is weakly Hölder continuous, there exists $\varepsilon > 0$ such that $s \mapsto P_G(\psi + s\upsilon)$ is real analytic on $(-\varepsilon, \varepsilon)$.*

We can use this to show that $s \mapsto P_G(\psi + s\upsilon)$ is analytic. However, to go from the Gurevich pressure to the usual pressure, we need a Variational Principle. Sarig's theory provides various conditions on potentials which yield a Variational Principle, but they are somewhat restrictive, and in particular for our case, are not satisfied by the potential $-t\log|Df|$. One aim of [FFY] is to weaken these conditions. There, the following theorem is proved.

Theorem 7. *If (W, S) be a transitive Markov shift and $\psi : W \to \mathbb{R}$ is a continuous function satisfying* (18)*, then $P_G(\psi) = P(\psi)$.*

We now apply Theorem 6 to the symbolic space induced by (Δ, f_Δ). In this space, the potential $(-t\log|Df| - S')_\Delta$ satisfies (18) and is continuous in the symbolic

metric. Theorem 6 implies that there is $\varepsilon' > 0$ such that $s \mapsto P_G(\psi_\Delta + s\upsilon_\Delta)$ is analytic on $(-\varepsilon', \varepsilon')$. Thus, by Theorem 7, $s \mapsto P(\psi_\Delta + s\upsilon_\Delta)$ is also analytic on $(-\varepsilon', \varepsilon')$.

All f_Δ-invariant probability measures ν have positive Lyapunov exponents. This is because the induced map (X, F) (which is isomorphic to the first return map to Δ_0) is uniformly expanding and the Ergodic Theorem gives

$$\lambda(\nu) := \int \log |Df_\Delta| \; d\nu = \nu(\Delta_0) \int \log |DF_\Delta| \; d\nu \geqslant \nu(\Delta_0) \inf_x \log |DF(x)| > 0.$$

Therefore $P(\psi_\Delta + s\upsilon_\Delta) = P_+(\psi_\Delta + s\upsilon_\Delta)$ for $s \in (-\varepsilon', \varepsilon')$.

Since the inducing scheme (X, F) is obtained from both (I, f) and (Δ, f_Δ) with the same inducing time $\tau = \tau_\Delta$, Lemma 5 implies that

$$h_{\mu_\Delta}(f_\Delta) = \left(\int \tau d\mu_F\right)^{-1} h_{\mu_F}(F) = h_\mu(f)$$

and

$$\mu_\Delta(\varphi_\Delta) = \left(\int \tau d\mu_F\right)^{-1} \mu_F(\Phi) = \mu(\varphi),$$

whenever μ_Δ and μ_F are the induced measures of μ to (Δ, f_Δ) and (X, F) respectively, and φ is any potential. Thus the free energy of μ and the lifted version μ_Δ are the same. This implies that $s \mapsto P_G(\psi + s\upsilon)$ is analytic on $(-\varepsilon', \varepsilon')$ if the definition of pressure involved only those measures which lift to Δ. Moreover, $P_+(\psi_\Delta + s\upsilon_\Delta) \leqslant P_+(\psi + s\upsilon)$ for $s \in (-\varepsilon', \varepsilon')$.

It remains to prove that there exists $\varepsilon > 0$ so that for all $s \in (-\varepsilon, \varepsilon)$, $P_+(\psi_\Delta + s\upsilon_\Delta) \geqslant P_+(\psi + s\upsilon)$. The issue is that in principle there might be measures which have high free energy but do not lift to Δ. We show how Lemma 10 implies that this is impossible, thus completing the theorem. Since by assumption $\sup_{\mu \in \mathcal{M}_+} |\int \upsilon \; d\mu| < \infty$, $P_+(\psi + \varepsilon\upsilon) \to P_+(\psi) = 0$ as $\varepsilon \to 0$. Therefore there exists $0 < \varepsilon < \varepsilon'$ so that for any $s \in (-\varepsilon, \varepsilon)$, we have $P_+(\psi + s\upsilon) > -\frac{\varepsilon_0}{2}$. Hence for all $s \in (-\varepsilon, \varepsilon)$, if a measure μ has $h_\mu(f) + \int \psi + s\upsilon \; d\mu > P_+(\psi + s\upsilon) - \frac{\varepsilon_0}{2}$ then Lemma 10 implies $\hat{\mu}(\hat{X}) > 0$. Hence $P_+(\psi_\Delta + s\upsilon_\Delta) \geqslant P_+(\psi + s\upsilon)$. Therefore $P_+(\psi_\Delta + s\upsilon_\Delta) = P_+(\psi + s\upsilon)$, and the analyticity of $s \mapsto P_+(\psi + s\upsilon)$ on $(-\varepsilon, \varepsilon)$ follows. □

It would be a further step to say that $t \mapsto \mu_{\varphi_t}$ is analytic (where μ_{φ_t} indicates the equilibrium state of φ_t). Using the weak topology we can ask whether $t \mapsto \int g \; d\mu_{\varphi_t}$ is analytic for any fixed continuous function g. We do have the following corollary:

Corollary 2. *In the setting of Theorems 1 and 2, let (X, F, τ) be any inducing scheme as in Section 3. Fix $s \in (t_1, 1)$ or s in a small neighbourhood of 1, according to whether (1) or (3) holds. Take $\psi_t = \varphi_t - P_+(\varphi_s)$ for $\varphi_t = -t \log |Df|$, and let Φ_t the induced potential. Then the function $t \mapsto \int_X \tau d\mu_{\Psi_t}$ is analytic for t sufficiently close to s, where μ_{Ψ_t} denotes the equilibrium state of Ψ_t.*

Proof. We know that $t \mapsto P_+(\psi_t)$ and $t \mapsto P(\Psi_t)$ are analytic. By Lemma 5, $P(\Psi_t) = (\int \tau d\mu_{\Psi_t}) \, P_+(\varphi_t)$, so analyticity of $t \mapsto \int \tau d\mu_{\Psi_t}$ follows. □

7. Concerning the Hypotheses of Theorems 1 and 2

In this section, we argue that the hypotheses of Theorems 1 and 2 cannot easily be relaxed. We also discuss some consequences of our proofs.

The set $\mathcal{M}_+$: The question how large the set $\mathcal{M}_+$ is in comparison to $\mathcal{M}_{erg}$ is answered by Hofbauer and Keller [HK3] in certain contexts. For unimodal maps, they prove that any measure $\mu \in \mathcal{M}_{erg} \setminus \mathcal{M}_+$ has entropy 0 and belongs to the convex hull of the set of weak accumulation points of $\{\frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k(c)}\}_{n \in \mathbb{N}}$, where $\delta_{f^k(c)}$ indicates the Dirac measure at the k-th image of the critical point. If we restrict to the potential $\varphi_t = -t \log |Df|$ at $t = 1$, then the following examples can be given:

- If f has a neutral fixed point, then the Dirac measure at this fixed point is an equilibrium state.
- There is a quadratic map without equilibrium measure for φ_1, see [BK]. In this case, the summability condition (12) fails.
- For maps such as the Fibonacci map (which satisfies (1) for $\ell = 2$), there is only one measure in $\mathcal{M}_{erg} \setminus \mathcal{M}_+$, namely the unique invariant probability measure $\mu_{\omega(c)}$ supported on the critical omega-limit set $\omega(c)$. This gives rise to a phase transition for the pressure function $t \mapsto P(\varphi_t)$ at $t = 1$. The quadratic Fibonacci map has two equilibrium states for φ_1: an absolutely continuous probability measure and $\mu_{\omega(c)}$.

  Moreover, there is a sequence of periodic points p_n with Lyapunov exponents $\lambda(p_n) \searrow 0$ as $n \to \infty$, see [NS]. The equidistributions on $\mathrm{orb}(p_n)$ belong to $\mathcal{M}_+$, which shows that $P_+(\varphi_t) = 0$ for $t \geqslant 1$, but $\mathcal{M}_+$ contains no equilibrium states if $t > 1$. See [BK] for more information on the phase transition.
- It is also possible that $\mathcal{M}_{erg} \setminus \mathcal{M}_+$ contains several equilibrium states, all supported on $\omega(c)$. In [B3] an example is given where $\omega(c)$ supports at least two ergodic measures, while there is also an acip, as follows from [B2, Theorem A (c)].

Differentiability of the map f: A $C^{1+\varepsilon}$ assumption is necessary in order to use the result that $\lambda(\mu) > 0$ implies liftability. This result, proved in [K1], relies on the property that μ-typical points have nondegenerate unstable manifolds, see [L]. If f is only piecewise continuous, this property as well as liftability no longer hold; this is illustrated by an example due to Raith [Ra], see the left-hand graph in Figure 1. This is piecewise continuous map f with slope 2, having a zero-dimensional set H on which f is semiconjugate to a circle rotation. The unique f-invariant measure μ of (H, f) has $\lambda(\mu) = \log 2 > 0$, but cannot be lifted to the Hofbauer tower, described in Section 3. This follows since it can be shown that for each $x \in H$ and $\hat{x} \in \pi^{-1}(x)$, $\hat{f}^n(\hat{x})$ belongs to a domain $D_n \in \mathcal{D}$ and $\lim_{n\to\infty} |D_n| \to 0$. As shown in the graph on the right of Figure 1, is easy to adjust this example into a continuous map with slope ± 2, but this map is not differentiable at the turning points. Another part where C^2 differentiability is used is Mañé's Theorem in the proof of Proposition 3.

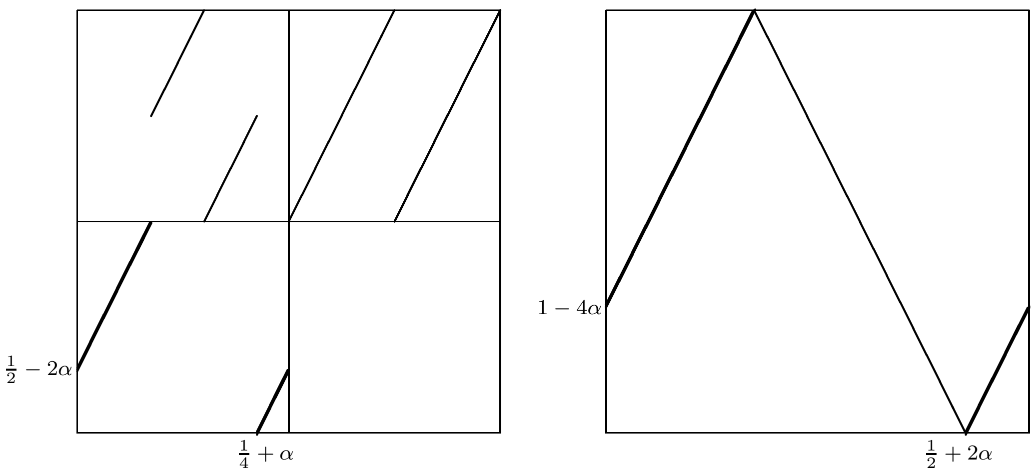

FIGURE 1. **Left**: Raith's example. For specific choices of α, the points whose orbits stay in the domains of branches 1 and 4 (bold lines) for ever form a zero-dimensional Cantor set H on which f is semi-conjugate to a circle rotation.
Right: Rescaling the left bottom square and inserting a new branch gives a continuous example. Again the set of points whose orbits stay in the domains branches 1 and 3 (bold lines) for ever form a zero-dimensional Cantor set H on which f is semi-conjugate to a circle rotation.

Measures with supp($\boldsymbol{\mu}$) $\subset$ orb(Crit): Makarov and Smirnov [MSm1, MSm2] discuss specific polynomials f on the complex plane for which there is a phase transition for the potential $\varphi_t = -t \log |Df|$ at some $t < 0$, and consequently these example would contradict our main theorem. The reason for this is that the Julia set $J(f)$ has 'very exposed' fixed points on which the Dirac measures can become equilibrium states for t sufficiently small. In the interval setting this applies to the Chebyshev polynomials $f : [0,1] \to [0,1]$ of any degree $d \geqslant 2$. The set $\{0,1\}$ consists of the critically accessible points; each critical point is prefixed, and either (a) $0 = f(0) = f(1) = f^2(\text{Crit})$; or (b) $0 = f(0)$, $f(1) = 1$ and 0 and 1 are both critical values of critical points. The critical accessibility creates an obstruction in our strategy of finding an induced scheme in Section 3. Further results on phase transitions for $t > 1$ are given in [MSm3].

The Gibbs property: Although the equilibrium states obtained in $\mathcal{M}_+$ (i.e., for the original system) are positive on open sets, we cannot expect them to be Gibbs. First, if $\varphi = -\log |Df|$, then φ is unbounded near critical points, so it is impossible to have $e^{\varphi_n(x) - nP(\varphi)} \leqslant K\mu(\mathbf{C}_n[x])$ uniformly in x. But also if the number K is allowed to depend on x, measures cannot always satisfy this weaker form of the Gibbs property. For example, if $f(x) = ax(1-x)$ has an acip μ, and the potential

is $\varphi = -\log|Df|$, then the pressure $P(\varphi) = 0$ and it is well known that $\frac{d\mu}{dx} > \rho_0 > 0$ on a neighbourhood of c. Suppose by contradiction that for each $x \notin \cup_{n\in\mathbb{Z}} f^n(c)$, there exists $K = K(x)$ such that

$$\frac{1}{K} \leqslant \frac{\mu(\mathbf{C}_n[x])}{e^{\varphi_n(x)}} \leqslant K \text{ for each } n \geqslant 0.$$

Now μ-a.e. x has an orbit accumulating on c, so almost surely there exists n such $|f^n(x) - c| < \frac{1}{4K^2}$. But then

$$\mu(\mathbf{C}_{n+1}[x]) \geqslant \frac{1}{K} e^{\varphi_{n+1}(x)} = \frac{1}{K} e^{\varphi_n(x)} \frac{1}{|Df^n(x)|} \geqslant \frac{1}{K^2} \mu(\mathbf{C}_n[x]) \frac{4K^2}{2} \geqslant 2\mu(\mathbf{C}_n[x]),$$

which contradicts that $\mathbf{C}_{n+1}[x] \subset \mathbf{C}_n[x]$. Thus μ cannot be a Gibbs measure.

In some cases, a weak Gibbs property can be proved. For example, it was shown in [BV] that for unimodal maps with critical order ℓ satisfying a summability condition, and every $\varepsilon > 0$, there exists $K = K(x)$ for Lebesgue a.e. x such that

$$\frac{1}{Kn^{3(\ell+1)}} \leqslant \frac{\mu_\varphi(\mathbf{C}_n[x])}{e^{\varphi_n(x)}} \leqslant Kn^{2(1+\varepsilon)}.$$

Appendix

In this appendix we give the two remaining proofs. The first is a lemma on the structure of the Hofbauer tower.

Proof of Lemma 1. We start with case (a), so Ω is a finite union of intervals. Let $x \in \Omega$ be any point with a dense orbit in Ω. Suppose that $(\mathcal{E}, \to)$ is a maximal primitive subgraph that is not closed, then for any $\hat{x} \in \pi^{-1}(x) \cap D_0$ for some $D_0 \in \mathcal{E}$, $\mathrm{orb}(\hat{x})$ leaves $\mathcal{E}$, i.e. $\hat{f}^k(\hat{x}) \notin \mathcal{E}$ for k sufficiently large. Indeed, since $\mathcal{E}$ is not closed, there is $D \in \mathcal{E}$ and $D' \notin \mathcal{E}$ such that $D \to D'$. There is an n-path $D_0 \to \cdots \to D$ for arbitrarily large n, corresponding to sets $\hat{\mathbf{C}}_n \in \hat{\mathcal{P}}_n$. Each $\hat{\mathbf{C}}_n$ has an $n+1$-subcylinder $\hat{\mathbf{C}}_{n+1}$ corresponding to the $n+1$-path $D_0 \to \cdots \to D \to D'$. For n sufficiently large, $\hat{\mathbf{C}}_{n+1}$ is compactly contained in D. Since $\mathrm{orb}(x)$ is dense in Ω, there is m such that $f^m(x) \in \pi(\hat{\mathbf{C}}_{n+1})$. Therefore $\hat{f}^m(\hat{x}) \in \pi^{-1} \circ \pi(\hat{\mathbf{C}}_{n+1})$ and $\hat{f}^{m+n+1}(\hat{x}) \in D''$ for some domain such that $\pi(D'') \subset \pi(D')$. Regardless of whether $D'' = D'$ or not, there is no path from D'' back into $\mathcal{E}$, because if there was, there would be a path from D' back into $\mathcal{E}$, contradicting maximality of $\mathcal{E}$.

Consequently, $\mathrm{orb}(\hat{x})$ will leave every maximal primitive subgraph that is not closed. If there is a closed primitive subgraph $(\mathcal{E}, \to)$, then it is unique, $\hat{f}^k(\hat{x}) \in \mathcal{E}$ for all sufficiently large k and necessarily $\pi(\cup_{D\in\mathcal{E}} D) \supset \Omega$. Let us also show that there is $\hat{y}$ with a dense orbit in $\mathcal{E}$. Fix $D_0 \in \mathcal{E}$ and let U_n be a countable base of $\sqcup_{D\in\mathcal{E}} D$. Each U_n intersects some D and U_n contains an r_n-cylinder $\hat{\mathbf{C}}_{r_n} \in \hat{\mathcal{P}}_n$ which itself is contained in D. Since $\mathcal{E}$ is primitive, there is a path $D_0 \to \cdots \to D$ of length l_n and another path $D \to \cdots \to D_0$ of length $l'_n \geqslant r_n$ such that if $\hat{z} \in D$ takes this path, then $\hat{z} \in \hat{\mathbf{C}}_{r_n}$. Let $p_n := l_n + l'_n$. Because $(\mathcal{E}, \to)$ is a Markov graph, for each $n \geqslant 1$ we have a cylinder $\hat{\mathbf{C}}_{p_n} \subset D_0$ such that $\hat{f}^{l_n}(\hat{\mathbf{C}}_{p_n}) \subset \hat{\mathbf{C}}_{r_n} \subset U_n$ and $\hat{f}^{p_n}(\hat{\mathbf{C}}_{p_n}) = D_0$.

Let $q_0 = 0$ and $q_n := \sum_{k=1}^n p_k$. Let $\hat{\mathbf{C}}_{q_1} = \hat{\mathbf{C}}_{p_1}$. By the Markov structure, we can pull back inductively to obtain a nested sequence of cylinder sets $\hat{\mathbf{C}}_{q_n} \subset \cdots \subset \hat{\mathbf{C}}_{q_1} \subset D_0$ with $\hat{f}^{q_n+l_{n+1}}(\hat{\mathbf{C}}_{q_{n+1}}) \subset U_{n+1}$ and $\hat{f}^{q_n}(\hat{\mathbf{C}}_{q_{n+1}}) = \hat{\mathbf{C}}_{p_{n+1}}$ for all $n \geqslant 0$. The point $\hat{y} \in \bigcap_n \hat{\mathbf{C}}_{q_n}$ has a dense orbit in $\mathcal{E}$. In this case the lemma is proved.

Alternatively, suppose that no closed primitive subgraph exists. Abbreviate $\hat{\Omega}_R := \pi^{-1}(\Omega) \cap \hat{I}_R$. If $\#(\text{orb}(\hat{x}) \cap \hat{\Omega}_R) = \infty$ for some R, then $\#(\text{orb}(\hat{x}) \cap D) = \infty$ for some $D \subset \hat{\Omega}_R$, and $\hat{f}^k(\hat{x})$ is in the non-empty maximal primitive subgraph containing D, for all sufficiently large k. The above argument shows that this subgraph is closed as well, so we would be in the previous case after all.

Therefore $\text{orb}(\hat{x})$ has a finite intersection with every compact subset of $\hat{I}$. We will show that this contradicts $\text{orb}(x)$ being dense in I, by showing that $\text{orb}(x)$ cannot accumulate on an orientation reversing fixed point p, leaving the (very similar) argument where p is orientation preserving and/or where p has a higher period to the reader.

Assume (for the moment) that all critical points are turning points (and not inflection points). Call ζ a *precritical point of order* k if $f^k(\zeta) \in \text{Crit}$ and $f^i(\zeta) \notin \text{Crit}$

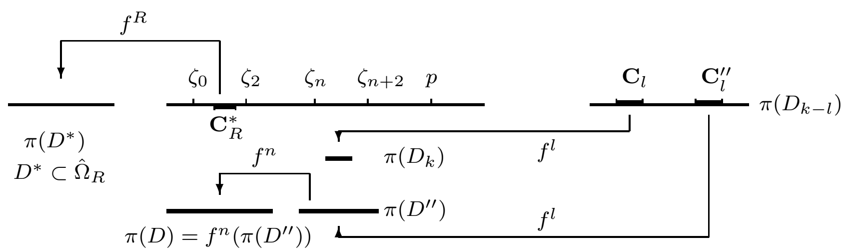

FIGURE 2. The π-images of domains $D = D_k$ and D', their positions with respect to ζ_n and a sketch how this leads to a path from D_{k-l} back into $\hat{\Omega}_R$.

for $i < k$. Let p be an orientation reversing fixed point and ζ_0 be a precritical point such that (ζ_0, p) contains no precritical point of lower order. Then there is a point $\zeta_1 \in f^{-1}(\zeta_0)$ at the other side of p with no precritical point of lower order in (p, ζ_1). Continue iterating backwards to find a sequence $\zeta_0 < \zeta_2 < \zeta_4 < \cdots < p < \cdots < \zeta_5 < \zeta_3 < \zeta_1$, such that (ζ_n, p) (or (p, ζ_{n+1})) contains no precritical point of lower order. Let R be such that (ζ_0, ζ_2) compactly contains an R-cylinder $\mathbf{C}^*_R$. It follows that if D is a domain such that $\pi(D) \supset (\zeta_0, \zeta_2)$, then there is an R-path from D leading to $D^* \subset \hat{\Omega}_R$, see Figure 2. To continue the argument, we need the following claim which is proved at the end of this proof.

Claim. *Take $\varepsilon := \min\{|c - c'| : c \neq c' \in \text{Crit}\}$, fix $l \geqslant 0$ and let J be any interval such that $|f^i(J)| < \varepsilon$ for all $i \leqslant l$. Then for any pair of l-cylinders $\mathbf{C}_l, \mathbf{C}'_l \subset J$, there is an l-cylinder $\mathbf{C}''_l$ in the convex hull of $\mathbf{C}_l$ and $\mathbf{C}'_l$ such that the images $f^l(\mathbf{C}_l), f^l(\mathbf{C}'_l) \subset f^l(\mathbf{C}''_l)$.*

Let D_k be the domain containing $\hat{f}^k(\hat{x})$. Recall that for every maximal primitive non-closed subgraph $\mathcal{E}$, $D_k \in \mathcal{E}$ for at most finitely many k. So let k_0 be such that D_{k_0} does not belong to any maximal primitive subgraph that intersects $\hat{\Omega}_R$. It follows that for each $k \geqslant k_0$, there is no path from D_k leading back into $\hat{\Omega}_R$. Furthermore, if $\limsup_k |D_k| \geqslant \varepsilon$, where ε is as in the claim, then for arbitrarily large k, there are paths D_k leading back into $\hat{\Omega}_R$. Therefore we can take k_0 so large that $|D_k| < \varepsilon$ for all $k \geqslant k_0$.

Assume by contradiction that $p \in \overline{\mathrm{orb}(x)}$. Then there are arbitrarily large n such that if $k = k(n)$ is the first integer such that $f^k(x) \in (\zeta_n, \zeta_{n+1})$, then $k > k_0$. Now if $\pi(D_k) \supset (\zeta_n, \zeta_{n+2})$, then there is an n-path from $D_k \to \cdots \to D$ where $\pi(D) \supset (\zeta_0, \zeta_2)$, and hence an $n+R$-path leading back into $\hat{\Omega}_R$ (as in Figure 2). This contradicts the definition of k_0.

Otherwise, i.e., if $\pi(D_k) \not\supset (\zeta_n, \zeta_{n+2})$, then the claim implies that there exist l and l-cylinders $\mathbf{C}_l, \mathbf{C}''_l \subset \pi(D_{k-l})$ such that $f^l(\mathbf{C}_l) = \pi(D_k)$ while D'' is such that $\pi(D'') = f^l(\mathbf{C}''_l) \supset \pi(D_k)$ and $\pi(D'') \supset (\zeta_n, \zeta_{n+2})$, see Figure 2. Take l minimal with this property. As before, this gives an $l+n+R$-path leading from D_{k-l} to $\hat{\Omega}_R$. If $k - l > k_0$, then we have a contradiction again with the choice of k_0. However, we can repeat the argument for infinitely many n, and hence infinitely many k. If D_{k-l} has been used for one value of k, then at least one domain in $\hat{f}(D_{k-l})$ is the starting domain of a path leading into $\hat{\Omega}_R$. Minimality of l implies that the same D_{k-l} no longer serves for the next value of k. This proves that for n sufficiently large, $k - l > k_0$, and this contradicts the choice of k_0, proving the lemma.

Finally, if there are critical inflection points, then we can repeat the argument with a branch partition and Hofbauer tower that disregards the inflection points. Indeed, the above arguments made use only of the topological structure of f, so whether $f|_{\mathbf{C}_1}$ is diffeomorphic or only homeomorphic on $\mathbf{C}_1 \in \mathcal{P}_1$ makes no difference.

Proof of the Claim. Let J be an interval such that $|J| < \varepsilon$. We argue by induction. For $l = 1$, the claim is true, since J can contain at most one 1-cylinder. Suppose now the claim holds for all integers $< l$ and $|f^i(J)| < \varepsilon$ for all $i \leqslant l-1$. Let $\mathbf{C}_l, \mathbf{C}'_l \subset J$ be l-cylinders, contained in $l-1$-cylinders $\mathbf{C}_{l-1}, \mathbf{C}'_{l-1}$. By induction, we can find an $l-1$-cylinder $\mathbf{C}''_{l-1}$ in the convex hull $[\mathbf{C}_{l-1}, \mathbf{C}'_{l-1}]$ such that $f^{l-1}(\mathbf{C}_{l-1}), f^{l-1}(\mathbf{C}'_{l-1}) \subset f^{l-1}(\mathbf{C}''_{l-1})$. If $\mathrm{Crit} \cap f^{l-1}(\mathbf{C}''_{l-1}) = \emptyset$ then $\mathbf{C}''_{l-1}$ is also an l-cylinder and $f^l(\mathbf{C}_l), f^l(\mathbf{C}'_l) \subset f^l(\mathbf{C}''_{l-1})$, proving the induction hypothesis for l. Otherwise, by definition of ε, $f^{l-1}(\mathbf{C}''_{l-1})$ contains a single critical point, and the f^l-image of one l-subcylinder of $\mathbf{C}''_{l-1}$ contains the f^l-image of the other. It is easy to see that this l-subcylinder satisfies the claim. □

This completes the proof of the claim and hence of part (a) of Lemma 1. Part (b) deals with renormalisable maps, so assume that $J \neq I$ is a p-periodic interval which is minimal in the sense that no proper subinterval of J has period p. We claim that J is associated with an absorbing subgraph $(\mathcal{E}_{\mathrm{absorb}}, \to)$ of $(\mathcal{D}, \to)$. Indeed, by minimality of J, $f^p : J \to J$ is onto, and for any $x \in \mathrm{orb}(J)$ and $n \geqslant 0$, there is

$x_n \in \mathrm{orb}(J)$ such that $f^n(x_n) = x$. Let $\hat{J} = \cap_k \hat{f}^k(\pi^{-1}(\mathrm{orb}(J)))$. This set has the following properties:

- $\hat{J} \neq \emptyset$: Since J contains an (interior) p-periodic point, it lifts to a p-periodic point in $\hat{J}$.
- If $\hat{x} \in \hat{J}$ and $D \in \mathcal{D}$ is the domain containing $\hat{x}$, then $D \subset \hat{J}$. This follows from the Markov property. Let $x = \pi(\hat{x})$, take $x_n \in \mathrm{orb}(J)$ as above and $\hat{x}_n \in \pi^{-1}(\mathrm{orb}(J))$ such that $\hat{f}^n(\hat{x}_n) = \hat{x}$. For $\hat{y} \in D$ arbitrary, we can find $\hat{y}_n \in \hat{Z}_n[\hat{x}_n]$ such that $\hat{f}^n(\hat{y}_n) = \hat{y}$. Since this holds for all $n \in \mathbb{N}$, $\hat{y} \in \hat{J}$.
- $\hat{J}$ is $\hat{f}$-invariant. This is immediate from the f-invariance of $\mathrm{orb}(J)$ and the definition of $\hat{J}$.

Take $\mathcal{E}_{\mathrm{absorb}} := \{D \in \mathcal{D} : D \cap \hat{J} \neq \emptyset\}$. Then the $\hat{f}$-invariance of $\hat{J}$ implies that $(\mathcal{E}_{\mathrm{absorb}}, \to)$ is indeed absorbing. Now apply part (a) to the subgraph $(\mathcal{D} \setminus \mathcal{E}_{\mathrm{absorb}}, \to)$ to find the required (non-closed) primitive subgraph. $\square$

The next proof shows that measures of positive entropy must lift to cover a large portion of the Hofbauer tower.

Proof of Lemma 3. Liftability of μ was shown by Keller [K1], so it remains to show that $\hat{\mu}(\hat{I}_R) > \eta$ uniformly over all measures with $h_\mu(f) \geqslant \varepsilon$.

Fix $R \in \mathbb{N}$ and $\delta > 0$ such that $(\delta + \frac{2}{R})\log(1 + \#\mathrm{Crit}) < \varepsilon/2$. Let $\mathcal{P}_n^u$ be the collection of n-cylinders such that $\frac{1}{n}\#\{k < n : \hat{f}^k \circ i(\mathbf{C}_n) \subset \hat{I}_R\} < \delta$, where as before $i^{-1} = \pi|_{D_0}$, and let $\mathcal{P}_n^l$ be the remaining n-cylinders.

If $\hat{\mu}(\hat{I}_R)$ is small, then $\mu(\cup_{\mathbf{C}_n \in \mathcal{P}_n^l} \mathbf{C}_n)$ is small as well. Hence, if the lemma was false, then for any $\eta > 0$ we could find a measure μ with $h_\mu(f) \geqslant \varepsilon$ and $\mu(\cup_{\mathbf{C}_n \in \mathcal{P}_n^l} \mathbf{C}_n) < \frac{\varepsilon}{2\log(1+\#\mathrm{Crit})}$. So assume by contradiction that there is such a measure μ.

If $D \in \mathcal{D}$ is any domain outside $\hat{I}_R$, then only the two outermost cylinder sets in $\mathcal{P}_R \cap D$ can map under $\hat{f}^R$ to domains of level $> R$. The $\hat{f}^R$-images of the other cylinder sets J' have both endpoints of level $\leqslant R$, so they have $\mathrm{level}(\hat{f}^R(J')) \leqslant R$. Repeating this argument for $\hat{f}^R(J')$ of those outermost cylinder sets, we can derive that for infinitely many n:

$$\lambda_u^n := \#\mathcal{P}_n^u \leqslant (1 + \#\mathrm{Crit})^{\delta n}(1 + \#\mathrm{Crit})^{(1-\delta)2n/R} \text{ and } \lambda_l^n := \#\mathcal{P}_n^l \leqslant (1 + \#\mathrm{Crit})^n,$$

so $\log \lambda_u \leqslant (\delta + \frac{2}{R})\log(1 + \#\mathrm{Crit}) < \varepsilon/2$ and $\log \lambda_l \leqslant \log(1 + \#\mathrm{Crit})$. For any finite set of nonnegative numbers a_k such that $\sum_k a_k = a \leqslant 1$, Jensen's inequality gives $-\sum_k a_k \log a_k \leqslant a \log \#\{a_k\}$. Since the branch partition $\mathcal{P}$ is assumed to generate

the Borel σ-algebra, the entropy of μ can be computed as

$$
\begin{aligned}
h_\mu(f) &= \inf_n -\frac{1}{n} \sum_{\mathbf{C}_n \in \mathcal{P}_n} \mu(\mathbf{C}_n) \log \mu(\mathbf{C}_n) \\
&= \inf_n -\frac{1}{n} \left(\sum_{\mathbf{C}_n \in \mathcal{P}_n^l} \mu(\mathbf{C}_n) \log \mu(\mathbf{C}_n) + \sum_{\mathbf{C}_n \in \mathcal{P}_n^u} \mu(\mathbf{C}_n) \log \mu(\mathbf{C}_n) \right) \\
&\leqslant \inf_n \frac{1}{n} \left(\frac{\varepsilon}{2(1+\#\mathrm{Crit})} \log \lambda_l^n + \log \lambda_u^n \right) < \varepsilon.
\end{aligned}
$$

This contradiction establishes the required $\eta > 0$.

Now to prove the second statement, for each $D \subset \hat{I}_R$, we can find $\kappa_D > 0$ such that if $\hat{x} \in D$ and $d(\hat{x}, \partial D) < \kappa_D$, then $\hat{f}^k(\hat{x}) \notin \hat{I}_R$ for $R < k \leqslant 3R/\eta$. Obviously the set $\hat{E} := \cup_{D \subset \hat{I}_R} \{\hat{x} \in D : d(\hat{x}, D) > \kappa_D\}$ is compactly contained in $\hat{I}_R$. If $\hat{x}$ is a typical point for $\hat{\mu}$, then the relative time of $\mathrm{orb}(\hat{x})$ spent outside $\hat{I}_R$ is at least $\hat{\mu}(\hat{I}_R \setminus \hat{E})(\frac{3}{\eta} - 1) \leqslant 1$, so $\hat{\mu}(\hat{I}_R \setminus \hat{E}) < \eta/2$, whence $\hat{\mu}(\hat{E}) > \eta/2$. □

References

[Ab] L. M. Abramov, *The entropy of a derived automorphism,* Dokl. Akad. Nauk SSSR **128** (1959) 647–650.

[Ba] V. Baladi, *Positive transfer operators and decay of correlations*, Advanced Series in Nonlinear Dynamics, **16** World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[Bo] R. Bowen, *Symbolic dynamics for hyperbolic systems,* Recent advances in topological dynamics (Proc. Conf., Yale Univ., New Haven, Conn., 1972; in honor of Gustav Arnold Hedlund), pp. 51–58. Lecture Notes in Math. **318** Springer, Berlin, 1975.

[B1] H. Bruin, *Induced maps, Markov extensions and invariant measures in one–dimensional dynamics,* Commun. Math. Phys., **168** (1995) 571–580.

[B2] H. Bruin, *Topological conditions for the existence of absorbing Cantor sets,* Trans. Amer. Math. Soc. **350** (1998) 2229–2263.

[B3] H. Bruin, *Minimal Cantor systems and unimodal maps,* J. Difference Equ. Appl. **9** (2003) 305–318.

[BK] H. Bruin, G. Keller, *Equilibrium states for S–unimodal maps,* Ergodic Theory Dynam. Systems **18** (1998) 765–789.

[BLS] H. Bruin, S. Luzzatto, S. van Strien, *Decay of correlations in one–dimensional dynamics*, Ann. Sci. Ec. Norm. Sup. **36** (2003) 621–646.

[BRSS] H. Bruin, J. Rivera-Letelier, W. Shen, S. van Strien, *Large derivatives, backward contraction and invariant densities for interval maps,* To appear in Invent. Math.

[BS] H. Bruin, S. van Strien, *Expansion of derivatives in one–dimensional dynamics*, Isr. J. Math. **137** (2003) 223–263.

[BT1] H. Bruin, M. Todd, *Markov Extensions and Lifting Measures for Complex Polynomials.* Ergodic Theory Dynam. Systems **27** (2007) 743–768.

[BT2] H. Bruin, M. Todd, *Equilibrium states for potentials with* $\sup\varphi - \inf\varphi < h_{top}(f)$*,* Preprint 2007, to appear in Commun. Math. Phys.

[BV] H. Bruin, S. Vaienti, *Return time statistics for unimodal maps,* Fund. Math. **176** (2003) 77–94.

[C] P. Collet, *Statistics of closest return times for some non uniformly hyperbolic systems,* Ergodic Theory Dynam. Systems **21** (2001) 401–420.

[FFY] D. Fiebig, U.-R. Fiebig, M. Yuri, *Pressure and equilibrium states for countable state Markov shifts,* Israel J. Math. **131** (2002) 221–257.

[H] F. Hofbauer, *The topological entropy of a transformation $x \mapsto ax(1-x)$,* Monath. Math. **90** (1980) 117–141.

[HK1] F. Hofbauer, G. Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations,* Math. Z. **180** (1982) 119–140.

[HK2] F. Hofbauer, G. Keller, *Equilibrium states for piecewise monotonic transformations,* Ergodic Theory Dynam. Systems **2** (1982) 23–43.

[HK3] F. Hofbauer, G. Keller, *Quadratic maps without asymptotic measure,* Commun. Math. Phys. **127** (1990) 319–337.

[HR] F. Hofbauer, P. Raith, *Topologically transitive subsets of piecewise monotonic maps which contain no periodic points,* Monatsh. Math. **107** (1989) 217–239.

[K1] G. Keller, *Lifting measures to Markov extensions,* Monatsh. Math. **108** (1989) 183–200.

[K2] G. Keller, *Equilibrium states in ergodic theory,* London Mathematical Society Student Texts, 42. Cambridge University Press, Cambridge, 1998.

[KN] G. Keller, T. Nowicki, *Spectral theory, zeta functions and the distribution of periodic points for Collet–Eckmann maps,* Commun. Math. Phys. **149** (1992) 31–69.

[Ko] O. S. Kozlovski, *Getting rid of the negative Schwarzian derivative condition,* Ann. of Math. **152** (2000) 743–762.

[L] F. Ledrappier, *Some properties of absolutely continuous invariant measures on an interval,* Ergodic Theory Dynam. Systems **1** (1981) 77–93.

[MSm1] N. Makarov, S. Smirnov, *Phase transition in subhyperbolic Julia sets,* Ergodic Theory Dynam. Systems **16** (1996) 125–157.

[MSm2] N. Makarov, S. Smirnov, *On "thermodynamics" of rational maps. I. Negative spectrum,* Commun. Math. Phys.**211** (2000) 705–743.

[MSm3] N. Makarov, S. Smirnov, *On thermodynamics of rational maps. II. Non-recurrent maps,* J. London Math. Soc. **67** (2003) 417–432.

[MU] D. Mauldin, M. Urbański, *Gibbs states on the symbolic space over an infinite alphabet,* Israel J. Math. **125** (2001) 93–130.

[MN] I. Melbourne, M. Nicol, *Almost sure invariance principle for nonuniformly hyperbolic systems,* Commun. Math. Phys. **260** (2005) 131–146.

[MSt] W. de Melo, S. van Strien, *One dimensional dynamics,* Ergebnisse Series **25**, Springer–Verlag, 1993.

[MSz] M. Misiurewicz and W. Szlenk, *Entropy of piecewise monotone mappings,* Studia Math. **67** (1980) 45–63.

[NS] T. Nowicki, D. Sands, *Non-uniform hyperbolicity and universal bounds for S-unimodal maps,* Invent. Math. **132** (1998) 633–680.

[PSe1] Y. Pesin, S. Senti *Equilibrium measures for some one dimensional maps,* Preprint 2004 (http://www.math.psu.edu/pesin/publications.html).

[PSe2] Y. Pesin, S. Senti, *Thermodynamical formalism associated with inducing schemes for one–dimensional maps,* Moscow Journal of Mathematics, **5:3** (2005).

[Pr] F. Przytycki, *Lyapunov characteristic exponents are nonnegative,* Proc. Amer. Math. Soc. **119** (1993) 309–317.

[Ra] P. Raith, *Hausdorff–dimension für stückweise monotone Abbildungen,* PhD. thesis (English), Universität Wien, 1987.

[Ru1] D. Ruelle, *An inequality for the entropy of differentiable maps,* Bol. Soc. Brasil. Mat. **9** (1978) 83–87.

[Ru2] D. Ruelle, *Thermodynamic formalism,* Addison Wesley, Reading MA, 1978.

[St.P] M. St.Pierre, *Zur Hausdorff-Dimension spezieller invarianter Maße für Collet-Eckmann Abbildungen,* Diplomarbeit, Erlangen (1994).

[Sa1] O. Sarig, *Thermodynamic formalism for Markov shifts,* PhD. thesis, Tel–Aviv, 2000.

[Sa2] O. Sarig, *Phase transitions for countable Markov shifts,* Commun. Math. Phys. **217** (2001) 555-577.

[Sa3] O. Sarig, *Existence of Gibbs measures for countable Markov shifts,* Proc. Amer. Math. Soc. **131** (2003) 1751–1758.

[Se] S. Senti, *Dimension of weakly expanding points for quadratic maps,* Bull. Soc. Math. Fr. **131** 399-420 (2003).

[Si] Y. Sinai, *Gibbs measures in ergodic theory,* Uspehi Mat. Nauk **27** (1972), no. 4 (166) 21–64.
[SV] S. van Strien, E. Vargas, *Real bounds, ergodicity and negative Schwarzian for multimodal maps,* J. Amer. Math. Soc. **17** (2004) 749–782.
[T] M. Todd, *Distortion bounds for $C^{2+\eta}$ unimodal maps,* Fund. Math. **193** (2007) 37–77.
[Y] L.-S. Young, *Recurrence times and rates of mixing,* Israel. J. Math. **110** (1999) 153–188.
[Z] R. Zweimüller, *Invariant measures for general(ized) induced transformations,* Proc. Amer. Math. Soc. **133** (2005) 2283–2295

Department of Mathematics
University of Surrey
Guildford, Surrey, GU2 7XH
UK
h.bruin@surrey.ac.uk
http://www.maths.surrey.ac.uk/

Department of Mathematics
University of Surrey
Guildford, Surrey, GU2 7XH
UK[3]
mtodd@fc.up.pt
http://www.fc.up.pt/pessoas/mtodd/

[3] **Current address:**
Departamento de Matemática Pura
Rua do Campo Alegre, 687
4169-007 Porto
Portugal